Studia Scientiarum Mathematicarum Hungarica

VOLUME 24
NUMBER 1
1989

AKADÉMIAI KIADÓ, BUDAPEST

// STUDIA SCIENTIARUM MATHEMATICARUM HUNGARICA

A QUARTERLY OF THE HUNGARIAN ACADEMY OF SCIENCES

Studia Scientiarum Mathematicarum Hungarica publishes original papers on mathematics mainly in English, but also in German, French and Russian.

Studia Scientiarum Mathematicarum Hungarica is published in yearly volumes of four issues (mostly double numbers published semiannually) by

AKADÉMIAI KIADÓ
Publishing House of the Hungarian Academy of Sciences
H—1054 Budapest, Alkotmány u. 21.

Manuscripts and editorial correspondence should be addressed to

J. Merza
Managing Editor
P.O. Box 127
H—1364 Budapest

Subscription information

Orders should be addressed to

KULTURA Foreign Trading Company
P.O. Box 149
H—1389 Budapest

or to its representatives abroad.

Studia Scientiarum Mathematicarum Hungarica

Volume 24

Akadémiai Kiadó, Budapest
1989

STUDIA SCIENTIARUM MATHEMATICARUM HUNGARICA

Volume 24

CONTENTS

89-1073 — Szegedi Nyomda — Felelős vezető: Kónya Antal mb. igazgató

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 1—9

COHOMOLOGY OF THE NILPOTENT SUBALGEBRAS OF CURRENT LIE ALGEBRAS

B. L. FEIGIN and A. FIALOWSKI[1]

Introduction

In this paper we compute the one and two dimensional cohomology spaces of the maximal nilpotent subalgebras of affine Lie algebras with coefficients in the adjoint representation. We also prove one of the possible analogues of the Bott—Kostant theorem for current Lie algebras. This article contains the details of the results announced in [4].

Let $\mathfrak{g}$ be a complex semisimple finite-dimensional Lie algebra, $\mathfrak{g}=\mathfrak{n}_-\oplus\mathfrak{h}\oplus\mathfrak{n}_+$ its Cartan decomposition, $\hat{\mathfrak{g}}=\mathfrak{g}\otimes\mathbf{C}[t, t^{-1}]$ the corresponding current Lie algebra, i.e., the Lie algebra of functions $S^1\to\mathfrak{g}$ having a finite Laurent expansion, with the bracket given by the formula $[f, g](x)=[f(x), g(x)]$, $f, g\in\hat{\mathfrak{g}}$, $x\in S^1$. Note that $\hat{\mathfrak{g}}$ admits the natural grading $\hat{\mathfrak{g}}=\oplus\hat{\mathfrak{g}}_m$, where $\hat{\mathfrak{g}}_m=\mathfrak{g}\otimes t^m$. Let us denote $(\mathfrak{n}_+\otimes 1)\oplus\oplus(\mathfrak{g}\otimes t)\oplus(\mathfrak{g}\otimes t^2)\oplus\dots$ by $\hat{\mathfrak{n}}_+$ and $\mathfrak{g}\otimes\mathbf{C}[t]$ by $\mathfrak{g}[t]$; $\hat{\mathfrak{n}}_+$ and $\mathfrak{g}[t]$ inherit the grading from $\hat{\mathfrak{g}}$. We shall identify $\mathfrak{g}$ with $\mathfrak{g}\otimes 1\subset\hat{\mathfrak{g}}$.

Recall that a current algebra is the quotient of an affine Lie algebra by its centre ([11]). Note that the main idea in the investigation of the cohomology of current algebras (as well as the other Kac—Moody algebras) is the analogy with the theory of finite-dimensional semisimple complex Lie algebras. In particular, $\hat{\mathfrak{n}}_+$ is a counterpart of the maximal nilpotent subalgebra of a finite-dimensional semisimple Lie algebra. So, we can use the well-known methods for computing the cohomology with the help of the Laplace operator ([12]), the Bernstein—Gelfand—Gelfand resolvent ([2]) etc. In [10] V. Kac proved that the cohomology space of $\hat{\mathfrak{n}}_+$ with trivial coefficients is in one to one correspondence with the group algebra of the affine Weyl group. As a consequence he obtained the Kac—MacDonald identities.

Another approach to the cohomology of current algebras uses ideas from the cohomology theory of the Lie algebra of tangent vector fields on a smooth manifold ([9]). We are going to use both methods.

In [13] Leger and Luks computed $H^2(\mathfrak{n}_+;\mathfrak{n}_+)$ (for another computation see [18]). They used the following method. The cohomology of $\mathfrak{n}_+$ with coefficients in an irreducible finite-dimensional representation V of $\mathfrak{g}$ is well-known. Namely, the Bott—Kostant Theorem (see, [12], [2]) asserts that $\dim H^i(\mathfrak{n}_+; V)$ is equal to the number of elements of length i in the Weyl group of $\mathfrak{g}$. In particular, we know

[1] The work was partly done during a fellowship of the Alexander von Humboldt-Stiftung at the Max-Planck-Institut für Mathematik, Bonn.

1980 *Mathematics Subject Classification.* Primary 17B56; Secondary 17B65.
Key words and phrases. Nilpotent Lie algebra, cohomology, spectral sequence. Kac—Moody algebra.

$H^{\bullet}(\mathfrak{n}_+; \mathfrak{g})$, where $\mathfrak{g}$ is the adjoint representation. Consider now the exact sequences of $\mathfrak{n}_+$-modules:

$$0 \to \mathfrak{n}_+ \to \mathfrak{g} \to \mathfrak{g}/\mathfrak{n}_+ \to 0, \qquad 0 \to \mathfrak{h} \to \mathfrak{g}/\mathfrak{n}_+ \to \mathfrak{n}_+^* \to 0.$$

Here $(\mathfrak{g}/\mathfrak{n}_+)/\mathfrak{h}$ can be identified with $\mathfrak{n}_+^*$ by means of the Killing form. These sequences allow us to reduce the computation of $H^2(\mathfrak{n}_+; \mathfrak{n}_+)$ to that of $H^1(\mathfrak{n}_+; \mathfrak{n}_+^*)$, and this space can be determined directly. In this paper we compute $H^i(\hat{\mathfrak{n}}_+; \hat{\mathfrak{n}}_+)$ for $i=1, 2$, generalizing the method in [13]. Another approach to affine algebras is contained in [7].

The cohomology of current algebras is similar to the cohomology of the Lie algebra of vector fields on the circle. Our method is illustrated on this Lie algebra of vector fields in [5].

In Section 1 we prove a Theorem, analogous to the Bott—Kostant Theorem, while in Section 2 we calculate $H^i(\hat{\mathfrak{n}}_+; \hat{\mathfrak{n}}_+)$ for $i=1, 2$.

The authors are greatful to Dmitrij Fuks and George Leger for their useful comments.

1. Computation of $H^{\bullet}(\hat{\mathfrak{n}}_+; \hat{\mathfrak{g}})$

The Bott—Kostant Theorem can be generalized to affine Lie algebras at least in two ways. The most direct generalization is the following one: if V is an irreducible representation of the current algebra with dominant highest weight, then $\dim H^i(\hat{\mathfrak{n}}_+; V)$ is equal to the number of elements of length i in the Weyl group. The proof is similar to that of the finite-dimensional case. The adjoint representation, however, is not a module of highest weight.

In this Section we give another generalization of the Bott—Kostant Theorem, namely we compute the cohomology of $\hat{\mathfrak{n}}_+$ with coefficients in modules similar to the adjoint module consisting of functions on the circle S^1 with values in the representation space of $\mathfrak{g}$.

Let V be a representation of $\mathfrak{g}$, A a $\mathbf{C}$-algebra and φ: $\mathbf{C}[t, t^{-1}] \to A$ a homomorphism. Define a representation of $\hat{\mathfrak{g}}$ in $V \otimes A$ by the formula

$$(x \otimes f)(v \otimes a) = x(v) \otimes \varphi(f)a, \quad x \in \mathfrak{g}, \quad v \in V, \quad f \in \mathbf{C}[t, t^{-1}], \quad a \in A.$$

We need two special cases: $A = \mathbf{C}[t, t^{-1}]$, φ is the identity map and $A = \mathbf{C}$, $\varphi(f) = f(1)$. In the first case denote the module $V \otimes A$ by $\hat{V}$ and in the second case by V_1. The elements of $\hat{V}$ are rational functions $\mathbf{C} \to V$, regular outside the origin. The mapping, sending a function $\mathbf{C} \to V$ to its value at 1, is a homomorphism $\hat{V} \to V_1$.

The space $\hat{V}$ is endowed with an obvious module structure over $\mathbf{C}[t, t^{-1}]$ and multiplication by an element of $\mathbf{C}[t, t^{-1}]$ is a $\mathfrak{g}$-endomorphism of the $\hat{\mathfrak{g}}$-module $\hat{V}$. Notice that $\hat{V}$ is a graded $\hat{\mathfrak{g}}$-module, $\hat{V} = \bigoplus_{i \in \mathbf{Z}} \hat{V}_i$ where $\hat{V}_i = V \otimes t^i$.

Now we are going to investigate the cohomology of $\hat{\mathfrak{n}}_+$ with coefficients in $\hat{V}$. Denote by $C^{\bullet}(\hat{\mathfrak{n}}_+; \hat{V})$ the cochain complex of $\hat{\mathfrak{n}}_+$ with coefficients in the $\hat{\mathfrak{n}}_+$-module $\hat{V}$. The complex $C^{\bullet}(\hat{\mathfrak{n}}_+; \hat{V})$ is graded by weights: $C^{\bullet}(\hat{\mathfrak{n}}_+; \hat{V}) = \bigoplus_{m \in \mathbf{Z}} C^{\bullet}_{(m)}(\hat{\mathfrak{n}}_+; \hat{V})$, where for the cochain $\varphi \in C^q_{(m)}(\hat{\mathfrak{n}}_+; \hat{V})$ the weight of $\varphi(e_{i_1}, \ldots, e_{i_q})$ is $m + i_1 + \ldots + i_q$ (i_k is the weight of e_{i_k}).

LEMMA 1. *For all m the complexes* $C^{\cdot}_{(m)}(\hat{\mathfrak{n}}_+; \hat{V})$ *are isomorphic to each other and to the complex* $C^{\cdot}(\hat{\mathfrak{n}}_+; V_1)$.

In fact, the composition of the embedding $C^{\cdot}_{(m)}(\hat{\mathfrak{n}}_+; \hat{V}) \to C^{\cdot}(\hat{\mathfrak{n}}_+; \hat{V})$ and of the mapping $C^{\cdot}(\hat{\mathfrak{n}}_+; \hat{V}) \to C^{\cdot}(\hat{\mathfrak{n}}_+; V_1)$ induced by the homomorphism $\hat{V} \to V_1$ is an isomorphism.

From the above it follows that the space $H^{\cdot}(\hat{\mathfrak{g}}; \hat{V})$ is a $\mathbf{C}[t, t^{-1}]$-module. Lemma 1 can be reformulated as follows.

LEMMA 2. $H^{\cdot}(\hat{\mathfrak{g}}; \hat{V}) \cong \mathbf{C}[t, t^{-1}] \otimes_{\mathbf{C}} H^{\cdot}(\hat{\mathfrak{g}}; V_1)$.

Let us now deal with the computation of $H^{\cdot}(\hat{\mathfrak{n}}_+; V_1)$. The Lie algebra $\mathfrak{n}_+$ is embedded into $\mathfrak{g}[t]$, V_1 is naturally endowed with a $\mathfrak{g}[t]$-module structure, consequently the homomorphism

$$\nu\colon H^{\cdot}(\mathfrak{g}[t], \mathfrak{g}; V_1) \to H^{\cdot}(\mathfrak{g}[t]; V_1) \to H^{\cdot}(\hat{\mathfrak{n}}_+; V_1)$$

is defined. Let τ be the homomorphism

$$H^{\cdot}(\hat{\mathfrak{n}}_+) \otimes H^{\cdot}(\mathfrak{g}[t], \mathfrak{g}; V_1) \to H^{\cdot}(\hat{\mathfrak{n}}_+; V_1),$$

sending $u \otimes v$ to the cohomology class $u\nu(v)$.

PROPOSITION 1. *If V is a finite-dimensional representation of* $\mathfrak{g}$ *then* τ *is an isomorphism.*

The proof will be given below. Proposition 1 and Lemma 2 imply the basic result of this Section.

THEOREM 1. $H^i(\hat{\mathfrak{n}}_+; \hat{\mathfrak{g}}) \cong \mathbf{C}[t, t^{-1}] \otimes_{\mathbf{C}} H^{i-1}(\hat{\mathfrak{n}}_+)$ *for any nonnegative integer i.*

Indeed, set $V = \mathfrak{g}$. The space $H^i(\hat{\mathfrak{n}}_+, \hat{\mathfrak{g}})$ is a $\mathbf{C}[t, t^{-1}]$-module. It follows from Lemma 2 that $H^i(\hat{\mathfrak{n}}_+, \hat{\mathfrak{g}})$ is a free module of rank equal to $\dim H^i(\hat{\mathfrak{n}}_+, V_1)$. The cohomology of $\hat{\mathfrak{n}}_+$ with trivial coefficients is known (see for instance [8]). Using this result, it is not difficult to find the cohomology $H^{\cdot}$ of $\mathfrak{g} \otimes t \oplus \mathfrak{g} \otimes t^2 \oplus \dots$. We only need the following fact. The space $H^{\cdot}$ is a $\mathfrak{g}$-module, and $\operatorname{Hom}_{\mathfrak{g}}(\mathfrak{g}, H^{\cdot}) = 0$ if $i \neq 1$ and $\mathbf{C}$ if $i = 1$ (see [14]). As $H^i(\mathfrak{g}[t], \mathfrak{g}; V) \cong \operatorname{Hom}_{\mathfrak{g}}(V, H^i)$, this gives us that $H^i(\mathfrak{g}[t], \mathfrak{g}; \mathfrak{g}) = 0$ for $i \neq 1$ and is one-dimensional for $i = 1$. After this it is enough to apply Proposition 1 and we get $H^i(\hat{\mathfrak{n}}_+; V_1) = H^{i-1}(\hat{\mathfrak{n}}_+)$.

Let us prove now Proposition 1. Introduce two subalgebras of $\hat{\mathfrak{g}}$: $\bar{\mathfrak{g}} = (t-1)\mathfrak{g} \oplus (t-1)^2\mathfrak{g} \oplus \dots$ and $\bar{\mathfrak{n}} = \hat{\mathfrak{n}}_+ \cap \bar{\mathfrak{g}}$. Let G be a compact connected Lie group corresponding to a compact real form of $\mathfrak{g}$.

LEMMA 3.

$$H^{\cdot}(\bar{\mathfrak{n}}) \cong H^{\cdot}(\hat{\mathfrak{n}}_+) \otimes H^{\cdot}(\bar{\mathfrak{g}}) \otimes H^{\cdot}(\Omega G).$$

Here ΩG is the loop space of G.

PROOF. Since $\mathfrak{g}[t] = \hat{\mathfrak{n}}_+ + \bar{\mathfrak{g}}$ and $\bar{\mathfrak{n}} = \hat{\mathfrak{n}}_+ \cap \bar{\mathfrak{g}}$ we have $C^{\cdot}(\bar{\mathfrak{n}}) = C^{\cdot}(\hat{\mathfrak{n}}_+) \otimes_{C^{\cdot}(\mathfrak{g}[t])} C^{\cdot}(\bar{\mathfrak{g}})$. Here the tensor product is taken in the category of differential algebras. In such a situation there exists a spectral sequence (Eilenberg—Moore, see [16]), connecting the cohomology of these four differential algebras. This spectral sequence generalizes the Künneth formula [15]. It converges to $H^{\bullet}(\bar{\mathfrak{n}})$ and the second term is isomorphic to $\mathrm{Tor}_A\big(H^{\bullet}(\hat{\mathfrak{n}}_+), H^{\bullet}(\bar{\mathfrak{g}})\big)$, where $A=H^{\bullet}(\mathfrak{g}[t])$. We remark that $H^{\bullet}(\mathfrak{g}[t])\cong H^{\bullet}(\mathfrak{g})$ (see e.g. [3]) and $H^{\bullet}(\mathfrak{g})$ acts trivially on $H^{\bullet}(\hat{\mathfrak{n}}_+)$ and on $H^{\bullet}(\bar{\mathfrak{g}})$ (for $H^{\bullet}(\bar{\mathfrak{g}})$ this is trivial and for $H^{\bullet}(\hat{\mathfrak{n}}_+)$ this follows from the fact that the composition $H^{\bullet}(\mathfrak{g})\to H^{\bullet}(\mathfrak{n}_+)\to\to H^{\bullet}(\hat{\mathfrak{n}}_+)$ is trivial). It follows from this that the second term of the spectral sequence is isomorphic to $H^{\bullet}(\hat{\mathfrak{n}}_+)\otimes H^{\bullet}(\bar{\mathfrak{g}})\otimes \mathrm{Tor}_A(\mathbf{C}, \mathbf{C})$.

We will show now that $\mathrm{Tor}_A(\mathbf{C}, \mathbf{C})\cong H^{\bullet}(\Omega G)$. Indeed, the cohomology algebra of $\mathfrak{g}$ with trivial coefficients coincides with the cohomology algebra of G and by the Hopf Theorem it is commutative and free [17]. Using the computation of $\mathrm{Tor}_A(\mathbf{C}, \mathbf{C})$ for the free commutative algebra A (Proposition 7.3 from [16] and see also [1]) and the connection between the cohomology of G and ΩG, we obtain the isomorphism $\mathrm{Tor}_A(\mathbf{C}, \mathbf{C})\cong H^{\bullet}(\Omega G)$.

Now it can be shown that the spectral sequence degenerates (e.g. by indicating explicit cycles of $C^{\bullet}(\bar{\mathfrak{n}})$ which represent the generators of E_2, which we shall do at the end of this Section). Lemma 3 is proved.

The Lie algebra $\bar{\mathfrak{n}}$ is an ideal in $\hat{\mathfrak{n}}_+$ and $\hat{\mathfrak{n}}_+/\bar{\mathfrak{n}}\cong\mathfrak{g}$. In virtue of this, $\mathfrak{g}$ acts on $H^{\bullet}(\bar{\mathfrak{n}})$. The algebra $\mathfrak{g}$ acts trivially on $H^{\bullet}(\hat{\mathfrak{n}}_+)$ and on $H^{\bullet}(\Omega G)$, but on $H^{\bullet}(\bar{\mathfrak{g}})$ it acts in the standard way ($\bar{\mathfrak{g}}\cong\mathfrak{g}\otimes t\oplus\mathfrak{g}\otimes t^2\oplus\dots$ is an ideal of $\mathfrak{g}[t]$, $\mathfrak{g}[t]/\bar{\mathfrak{g}}\cong\mathfrak{g}$, so $\mathfrak{g}$ acts on $\bar{\mathfrak{g}}$ naturally and the action of $\mathfrak{g}$ on $H^{\bullet}(\bar{\mathfrak{g}})$ is semisimple).

Now to finish the proof of Theorem 1 let us consider the Serre—Hochschild spectral sequence, associated with $\hat{\mathfrak{n}}_+$, its ideal $\bar{\mathfrak{n}}$ and the module V_1, converging to $H^{\bullet}(\hat{\mathfrak{n}}_+; V_1)$. The algebra $\bar{\mathfrak{n}}$ acts on V_1 trivially. The second term of this spectral sequence is the following:

$$H^{\bullet}\big(\mathfrak{g};\ H^{\bullet}(\bar{\mathfrak{n}}, V_1)\big)\cong H^{\bullet}\big(\mathfrak{g};\ H^{\bullet}(\bar{\mathfrak{n}})\otimes V_1\big)\cong H^{\bullet}\big(\mathfrak{g};\ H^{\bullet}(\hat{\mathfrak{n}}_+)\otimes H^{\bullet}(\bar{\mathfrak{g}})\otimes H^{\bullet}(\Omega G)\otimes V_1\big)\cong$$

$$\cong H^{\bullet}(\hat{\mathfrak{n}}_+)\otimes H^{\bullet}(\Omega G)\otimes H^{\bullet}\big(\mathfrak{g};\ H^{\bullet}(\bar{\mathfrak{g}})\otimes V_1\big).$$

As $\mathfrak{g}$ is semisimple, $H^{\bullet}(\bar{\mathfrak{g}})\otimes V_1$ is the direct sum of finite-dimensional representations, i.e.

$$H^{\bullet}\big(\mathfrak{g};\ H^{\bullet}(\bar{\mathfrak{g}})\otimes V_1\big)\cong H^{\bullet}(\mathfrak{g})\otimes I$$

where I is the invariant space of $H^{\bullet}(\bar{\mathfrak{g}})\otimes V_1$ (see [8]). Note that $I\cong H^{\bullet}(\mathfrak{g}[t], \mathfrak{g}; V_1)$. The differentials in the above sequence act in the following way: they map the generators of the algebra $H^{\bullet}(\Omega G)$ into the generators of $H^{\bullet}(\mathfrak{g})$ and are trivial on $H^{\bullet}(\hat{\mathfrak{n}}_+)\otimes H^{\bullet}(\mathfrak{g}[t], \mathfrak{g}; V_1)$. It follows from this that the spectral sequence converges to $H^{\bullet}(\hat{\mathfrak{n}}_+)\otimes H^{\bullet}(\mathfrak{g}[t], \mathfrak{g}; V_1)$. Thus our spectral sequence is the product of $H^{\bullet}(\hat{\mathfrak{n}}_+)\otimes\otimes H^{\bullet}(\mathfrak{g}[t], \mathfrak{g}; V_1)$ with the spectral sequence of the Serre path fibration $EG\to G$; it follows from this that τ is an isomorphism.

Now we explain why the spectral sequence in the proof of Lemma 3 collapses. To define explicitly cycles of $C^{\bullet}(\bar{\mathfrak{n}})$, representing the generators of E_2 we apply the continuous cohomology theory. Let $\mathfrak{n}(0, 1)$ be the Lie algebra of infinitely differentiable functions f: $[0, 1]\to\mathfrak{g}$ such that $f(0)\in\mathfrak{n}$, $f(1)=0$. Denote by $C^{\bullet}_C(0, 1)$ the complex of cochains of $\mathfrak{n}(0, 1)$, continuous in the C^∞-topology. Let α be a generator of $H^{\bullet}(\mathfrak{g})$ and $\bar{\alpha}$ a cochain representing α. For $p\in[0, 1]$ denote by φ_p the homomorphism $\bar{\mathfrak{n}}\to\mathfrak{g}$, "the value at p": $\varphi_p((t-1)g_1, (t-1)^2g_2, \dots)=\sum_m (p-1)^m g_m$.

Let $\alpha_p=\varphi_p^*\bar{\alpha}$, $\alpha_p\in C^{\bullet}_C(0, 1)$. Choose $\bar{\alpha}$ in such a way that $\alpha_0=\alpha_1=0$. Let $p\neq 0, 1$;

then we can define the cochain $\frac{\partial\alpha}{\partial x}(p)$ where x is the coordinate on [0, 1]. It is shown in [3] that $\frac{\partial\alpha}{\partial x}(p)$ is a coboundary $\frac{\partial\alpha}{\partial x}(p)=\delta\omega(p)$ where δ is the differential in $C^{\cdot}_{C}(0, 1)$. Indeed, let $K_p(p\neq 0, 1)$ be the cochain complex of $\overline{\mathfrak{n}}$ with support at p. It is proved in the same paper that the cohomology of K_p is isomorphic to $H^{\cdot}(\mathfrak{g})$. Now, K_p is W_1-module, where W_1 is the Lie algebra of formal vector fields at the point p. But $H^{\cdot}(\mathfrak{g})$ is finite-dimensional and W_1 has no nontrivial finite-dimensional representations. We conclude that if $\omega\in K_p$ and $\delta v=0$ then $\frac{\partial}{\partial x}v$ is the differential of some other cocycle $\bar{v}\in K_p$.

This means that

$$\alpha_p-\alpha_q=\delta\int_p^q\omega(x)\,dx.$$

In particular, $\delta\int_0^1\omega(x)\,dx=0$. Suppose that $\alpha'=\int_0^1\omega(x)\,dx$. The cochain α' represents a nontrivial cohomology class of $\overline{\mathfrak{n}}$.

The Lie algebras $\hat{\mathfrak{n}}_+$ and $\bar{\mathfrak{g}}=\mathfrak{g}\otimes(t-1)\oplus\mathfrak{g}\otimes(t-1)^2\oplus\ldots$ are graded. Similarly the cochain complexes are also graded. Note that the cochain complex K_0 of $\mathfrak{n}(0, 1)$ with support in 0 is isomorphic to $\oplus C^{\cdot}_i(\hat{\mathfrak{n}}_+)$ and the cochain complex K_1 with support in 1 is isomorphic to $\oplus C^{\cdot}_i(\bar{\mathfrak{g}})$. It follows from this that the cohomology of K_0 and K_1 is isomorphic to $H^{\cdot}(\hat{\mathfrak{n}}_+)$ and $H^{\cdot}(\bar{\mathfrak{g}})$, respectively.

Recall that $H^{\cdot}(\mathfrak{g})$ is isomorphic to the free graded commutative algebra on generators $\xi_1, \xi_2, \ldots$, $\deg\xi_k=2k+1$. Using the above construction assign to each ξ_i a representative cocycle ξ_i'.

PROPOSITION 2. *The space $H^{\cdot}(\overline{\mathfrak{n}})$ is generated by the cohomology classes of cochains of form $u\wedge v\wedge P(\xi_1', \xi_2', \ldots)$, where $u\in K_0, v\in K_1$ are cocycles, corresponding to the elements of $H^{\cdot}(\hat{\mathfrak{n}}_+)$ and $H^{\cdot}(\bar{\mathfrak{g}})$ respectively and P is an arbitrary polynomial in generators $\xi_1', \xi_2', \ldots$.*

The proof of this Proposition follows from the construction above for continuous cohomology (a similar argument in a more difficult situation was used in [6]). In particular, we have an explicit construction of cochains, representing the generators of E_2 in the proof of Lemma 3, surviving till E_∞.

2. Computation of $H^i(\hat{\mathfrak{n}};\hat{\mathfrak{n}})$ for $i=1, 2$

Let us consider the next exact sequences:

$$0\to\hat{\mathfrak{n}}_+\to\hat{\mathfrak{g}}\to\hat{\mathfrak{g}}/\hat{\mathfrak{n}}_+\to 0;\qquad 0\to\mathfrak{h}\to\hat{\mathfrak{g}}/\hat{\mathfrak{n}}_+\to\hat{\mathfrak{n}}^*_+\to 0$$

($(\hat{\mathfrak{g}}/\hat{\mathfrak{n}})/\mathfrak{h}$ can be identified with $\hat{\mathfrak{n}}^*_+$ by means of the Killing form). Consider the induced exact cohomology sequences:

$$H^0(\hat{\mathfrak{n}}_+;\hat{\mathfrak{g}}/\hat{\mathfrak{n}}_+)\to H^1(\hat{\mathfrak{n}}_+;\hat{\mathfrak{n}}_+)\to H^1(\hat{\mathfrak{n}}_+;\hat{\mathfrak{g}})\to H^1(\hat{\mathfrak{n}}_+;\hat{\mathfrak{g}}/\hat{\mathfrak{n}}_+)\to$$
$$\to H^2(\hat{\mathfrak{n}}_+;\hat{\mathfrak{n}}_+)\to H^2(\hat{\mathfrak{n}}_+;\hat{\mathfrak{g}});$$
$$H^0(\hat{\mathfrak{n}}_+;\hat{\mathfrak{n}}^*_+)\to H^1(\hat{\mathfrak{n}}_+;\mathfrak{h})\to H^1(\hat{\mathfrak{n}}_+;\hat{\mathfrak{g}}/\hat{\mathfrak{n}}_+)\to H^1(\hat{\mathfrak{n}}_+;\hat{\mathfrak{n}}^*_+)\to H^2(\hat{\mathfrak{n}}_+;\mathfrak{h}).$$

The first sequence allows us to compute $H^1(\hat{\mathfrak{n}}_+; \hat{\mathfrak{n}}_+)$ at once. We will state the result, i.e. describe all the derivations of $\hat{\mathfrak{n}}_+$.

Each element $u \in \mathfrak{h}$ defines a cohomology class of $H^1(\hat{\mathfrak{n}}_+; \hat{\mathfrak{n}}_+)$ containing the cocycle $f \to uf$, $(uf)(t) = [u, f(t)]$, where f: $\mathbf{C} \to \mathfrak{g}$, $f(0) \in \mathfrak{n}_+$. Then to each vector field $tP\partial/\partial t$ where P is a polynomial in t, we assign the cocycle

$$\hat{\mathfrak{n}}_+ \to \hat{\mathfrak{n}}_+ : f(t) \mapsto tP\partial f(t)/\partial t, \quad f: \mathbf{C} \to \mathfrak{g}, \quad f(0) \in \mathfrak{n}_+.$$

THEOREM 2. *The mapping, sending the elements of* $\mathfrak{h}$ *and* $t\mathbf{C}[t]\partial/\partial t$ *to the cohomology classes of the cocycles constructed above gives an isomorphism* $h \oplus t\mathbf{C}[t]\partial/\partial t \cong$ $\cong H^1(\hat{\mathfrak{n}}_+; \hat{\mathfrak{n}}_+)$.

In other words, an arbitrary derivation of $\hat{\mathfrak{n}}$ is uniquely represented as $u + tP\,\partial/\partial t + q$ where $u \in \mathfrak{h}$, $P \in \mathbf{C}[t]$ and q is an inner derivation.

For the computation of $H^2(\hat{\mathfrak{n}}_+; \hat{\mathfrak{n}}_+)$ by a similar way, we have to know $H^1(\hat{\mathfrak{n}}_+; \hat{\mathfrak{g}}/\hat{\mathfrak{n}}_+)$. This space appears in the second exact sequence and to find it we have to know the isomorphism $H^1(\hat{\mathfrak{n}}_+; \hat{\mathfrak{n}}_+^*) \cong (H_1(\hat{\mathfrak{n}}_+; \hat{\mathfrak{n}}_+))^*$. For this we are going to use the next general construction (see Theorem 4.1 in [13]).

Let $\mathscr{L}$ be a Lie algebra, T a derivation of $\mathscr{L}$ acting in a semi-simple way and whose eigenvalues are positive. (These restrictions on T can be considerably weakened.) It is clear that such a derivation must be outer. It is easy to see that $\hat{\mathfrak{n}}_+$ has such a derivation. For instance, we can take $u + t\,\partial/\partial t$, where $u \in \mathfrak{h}$, $\langle \gamma, u \rangle \in \mathbf{R}$, $0 < \langle \gamma, u \rangle < \varepsilon$, γ is an arbitrary positive root, ε is a small positive number $\left(\text{we can take } \varepsilon < \dfrac{1}{2}\right)$.

Let $W(\mathscr{L})$ be the Weyl algebra, associated with $\mathscr{L}$. Recall that $W(\mathscr{L})$ is the standard complex of the differential Lie superalgebra $\bar{\mathscr{L}} = \mathscr{L}_0 \oplus \mathscr{L}_1$, $\mathscr{L}_0 \cong \mathscr{L}$, $\mathscr{L}_1$ as $\mathscr{L}_0$-module is the adjoint representation and $[x, x] = 0$ if $x \in \mathscr{L}_1$. The differential d acts as follows: $d(\mathscr{L}_0) = 0$, $d(\mathscr{L}_1) \to \mathscr{L}_0$ is an isomorphism of $\mathscr{L}$-modules. In other words, $W(\mathscr{L})$ is a differential graded algebra, spanned by $\mathscr{L}_0^*$ and $\mathscr{L}_1^*$, $\mathscr{L}_0^* \cong \mathscr{L}_1^* \cong \mathscr{L}^*$ where $\mathscr{L}_0^*$ has degree 1 and $\mathscr{L}_1^*$ has degree 2, φ: $\mathscr{L}_0^* \to \mathscr{L}_1^*$ is the canonical isomorphism. The differential is defined by the formula $\delta\beta = \delta_0\beta + \varphi(\beta)$ where $\delta_0\beta$ is the differential of β considered as an element of the standard cochain complex of $\mathscr{L}$. It is not difficult to show that the complex $W(\mathscr{L})$ is acyclic in positive dimensions. The next Lemma states even more.

In $W(\mathscr{L})$ we define a filtration: $W_i = \bigoplus_{j \geq i} \Lambda^*(\mathscr{L}_0^*) \otimes S^j(\mathscr{L}_1)^*$. Consider the corresponding spectral sequence E.

LEMMA 4. *The spectral sequence*

$$E_2^{p,q} = H^q(\mathscr{L}; S^{p/2}\mathscr{L}^*) \Rightarrow H(W(\mathscr{L}))$$

is trivial, beginning from E_3.

REMARK. For arbitrary algebra $\mathscr{L}$ this is of course not true, but in this Lemma we consider such $\mathscr{L}$ which satisfies a strong additional assumption: there exists such a semisimple derivation T: $\mathscr{L} \to \mathscr{L}$ for which all the eigenvalues are positive. Such a derivation can only exist in case of nilpotent algebras, and for them also not always. Let us consider in our case such a T.

PROOF of Lemma 4. The differential

$$\delta = d_2^{p,q}: H^q(\mathscr{L};\ S^{p/2}\mathscr{L}^*) \to H^{q-1}(\mathscr{L};\ S^{p/2+1}\mathscr{L}^*)$$

is defined on the cochain level by the formula $\left(\frac{p}{2}=r\right)$

$$[(\delta\varphi)(l_1, \dots, l_{q-1})](l_1', \dots, l_{r+1}') = \sum_{j=1}^{r+1} [\varphi(l_1, \dots, l_{q-1}, l_j')](l_1', \overset{\hat{j}}{\dots}, l_{r+1}').$$

Define the map D: $H^{q-1}(\mathscr{L}; S^{r+1}\mathscr{L}^*) \to H^q(\mathscr{L}; S^r\mathscr{L}^*)$ by the formula

$$[(D\varphi)(l_1, \dots, l_q)](l_1', \dots, l_r') = \sum_{i=1}^{p} (-1)^{p-i} [\varphi(l_1, \overset{\hat{i}}{\dots}, l_q)](Tl_i, l_1', \dots, l_r').$$

It is easy to check that the bracket $[D, \delta]$ coincides with the map, defined in $H^{\cdot}(\mathscr{L}; S^*\mathscr{L}^*)$ by T. This map has only positive eigenvalues and can be transposed with δ. Define D_0 as $\frac{1}{\lambda} D$ on the λ-eigenspace of T in $H^{\cdot}(\mathscr{L}; S^*\mathscr{L}^*)$. Then D_0 is a contracting homotopy in the complex $E_2 = \{H^{\cdot}(\mathscr{L}; S^*\mathscr{L}^*), \delta\}$ and this means that $E_3 = 0$.

This lemma implies in particular, that the sequence

$$0 \to H^2(\mathscr{L}) \to H^1(\mathscr{L}, \mathscr{L}^*) \to H^0(\mathscr{L}, S^2\mathscr{L}^*) \to 0$$

is exact. The arrows here are differentials in the second term of E. We remark that $H^0(\mathscr{L}, S^2\mathscr{L}^*)$ is exactly the space of invariant bilinear symmetric forms on $\mathscr{L}$.

THEOREM 3. *The space of invariant bilinear symmetric forms on $\hat{\mathfrak{n}}_+$ is the direct sum of the following two subspaces intersecting trivially.*

a) *The first space consists of the forms whose kernel contains* $[\hat{\mathfrak{n}}_+, \hat{\mathfrak{n}}_+]$. *This space is isomorphic to the space of quadratic forms on* $\hat{\mathfrak{n}}_+/[\hat{\mathfrak{n}}_+, \hat{\mathfrak{n}}_+]$, *i.e. has dimension* $(l+1)(l+2)/2$ *where l is the rank of* $\mathfrak{g}$.

b) *Let* $P(t^{-1})\partial/\partial t$ *be a vector field, where P is a polynomial without constant term. The second space consists of the forms*

$$(x, y) \mapsto \langle P(t^{-1})\partial x/\partial t, y\rangle + \langle P(t^{-1})\partial y/\partial t, x\rangle$$

where $x, y \in \hat{\mathfrak{n}}_+$, $\langle\ ,\ \rangle$ *is the Killing form on* $\hat{\mathfrak{g}}$.

PROOF. Let ω be an invariant bilinear symmetric form on $\hat{\mathfrak{n}}_+$. Let us assign to the quadratic form, associated with ω the mapping θ: $\hat{\mathfrak{n}}_+ \to \hat{\mathfrak{n}}_+^*$; here $\hat{\mathfrak{n}}_+ = \mathfrak{n}_+ \oplus \mathfrak{g}\otimes t \oplus \mathfrak{g}\otimes t^2 \oplus \dots$, $\hat{\mathfrak{n}}_+^* = \mathfrak{n}_+^* \oplus (\mathfrak{g}\otimes t)^* \oplus (\mathfrak{g}\otimes t^2)^* \oplus \dots$. Suppose that ω is homogeneous with respect to the grading by the weight of t; the $\hat{\mathfrak{n}}_+$-module $\hat{\mathfrak{n}}_+^*$ is filtrated by the submodules $\hat{\mathfrak{n}}_0^* = \mathfrak{n}_+^*$, $\hat{\mathfrak{n}}_1^* = \mathfrak{n}_+^* \oplus (\mathfrak{g}\otimes t)^*$ etc. Let i be the smallest number such that $\theta(\hat{\mathfrak{n}}_+) \subset \hat{\mathfrak{n}}_i^*$. If $i=0$ then ω lies in the first factor. The assertion that the kernel of this form contains $[\hat{\mathfrak{n}}_+, \hat{\mathfrak{n}}_+]$ follows from Theorem 5.1 in [13]. If $i=1$ then the image of the mapping $\mathfrak{n}_+ \to (\mathfrak{g}t)^*$ is either one-dimensional or coincides with $\mathfrak{n}_+ (\mathfrak{n}_+ \subset \mathfrak{g} \cong (\mathfrak{g}\otimes t)^*$, $\mathfrak{g}$ is identified with $(\mathfrak{g}\otimes t)^*$ by the Killing form). In the first case ω lies in the first factor and in the second case in the second one. These facts follow from the following simple Lemma.

LEMMA 5. $\dim \operatorname{Hom}_{\mathfrak{n}_+}(\mathfrak{n}_+, \mathfrak{g})=1+l,\ l>1.$

(This Lemma can be verified for instance by looking over all the simple Lie algebras.)

Further, by using Lemma 5, we get that if $i\geqq 2$, then the form belongs to the second space. This completes the proof of Theorem 3.

Now, using the exact sequence (3), we compute $H^1(\hat{\mathfrak{n}}_+; \hat{\mathfrak{n}}_+^*)$ and after this we can determine $H^2(\hat{\mathfrak{n}}_+; \hat{\mathfrak{n}}_+)$.

THEOREM 4. a) *The kernel of the natural mapping*

$$\varphi: H^1(\hat{\mathfrak{n}}_+)\times H^1(\hat{\mathfrak{n}}_+;\ \hat{\mathfrak{n}}_+) \to H^2(\hat{\mathfrak{n}}_+;\ \hat{\mathfrak{n}}_+)$$

is $l+1$-dimensional where l is the rank of $\mathfrak{g}$.

b) *If* $\operatorname{rank} \mathfrak{g}>1$, *then* $\dim \operatorname{coker} \varphi=l+1+p$ *where p is the number of positive roots of* $\mathfrak{g}$ *representable as the sum of two simple roots.*

The case $\mathfrak{g}=\mathrm{sl}(2, \mathbf{C})$ is not covered by this Theorem. This case is really ar exceptional one (see [7]).

REFERENCES

[1] AVRAMOV, L. L., On the Hopf algebra of a local ring, *Izv. Akad. Nauk. SSSR Ser. Mat.* **38** (1974), 253—277 (in Russian). *MR* **50** #2309.

[2] BERNSTEIN, J. N., GELFAND, I. M. and GELFAND, S. I., Differential operators on the base affine space and a study of $\mathfrak{g}$-modules, *Lie groups and their representations* (Proc. Summer School, Bolyai János Math. Soc., Budapest, 1971), Akadémiai Kiadó, Budapest; Halsted, New York, 1975. *MR* **58** #28285.

[3] FEIGIN, B. L., Cohomology of groups and of algebras of flows, *Uspehi Mat. Nauk* **35** (1980), no. 2, 225—226. *MR* **81h:** 17018.

[4] FEIGIN, B. L. and FIALOWSKI, A., The cohomology of nilpotent algebras of flows, *Dokl. Akad. Nauk SSSR* **271** (1983), no. 4, 813—816 (in Russian). *MR* **84k:** 17013.

[5] FEIGIN, B. L. and FIALOWSKI, A., Cohomology of the infinite dimensional Lie algebra L_1 with nontrivial coefficients, *Bull. Amer. Math. Soc. (N. S.)* **17** (1987), no. 2, 333—337.

[6] FEIGIN, B. L. and RETAKH, V. S., Cohomology of some Lie algebras and superalgebras of vector fields, *Uspekhi Mat. Nauk* **37** (1982), no. 2, 233—234. *MR* **83m:** 58083.

[7] FIALOWSKI, A., Deformations of nilpotent Kac—Moody algebras, *Studia Sci. Math. Hungar.* **19** (1984), 465—483.

[8] FUKS D. B., *Cohomology of infinite dimensional Lie algebras.* Moscow, Nauka, 1984 (in Russian); Consultants Bureau, New York, 1986 (in English).

[9] GELFAND, I. M. and FUKS, D. B., Cohomologies of the Lie algebra of tangent vector fields of a smooth manifold, *Funkcional. Anal. i Priložen.* **3** (1969), no. 3, 32—52 (in Russian). *MR* **41** #1067.

[10] KAC, V., Infinite-dimensional Lie algebras and the Dedekind η-function, *Funkcional Anal. i Priložen.* **8** (1974), no. 1, 77—78 (in Russian).

[11] KAC, V., *Infinite-dimensional Lie algebras.* An introduction, Progress in Mathematics, Vol. 44, Birkhäuser, Boston, 1983. *MR* **86h:** 17015.

[12] KOSTANT, B., Lie algebra cohomology and the generalized Borel—Weil Theorem, *Ann. of Math.* (2) **74** (1961), 329—387. *MR* **26** #265.

[13] LEGER, G. F. and LUKS, E. M., Cohomology of nilradicals of Borel subalgebras, *Trans. Amer. Math. Soc.* **195** (1974), 305—316. *MR* **51** #808.

[14] LEPOWSKY, J., Generalized Verma modules, loop space cohomology and MacDonald-type identities, *Ann. Sci. Ecole Norm. Sup.* (4) **12** (1979), 169—234. *MR* **81a:** 17004.

[15] MACLANE, S. *Homology,* Die Grundlehren der mathematischen Wissenschaften, Bd. 114, Academic Press, New York; Springer, Berlin—Göttingen—Heidelberg, 1963. *MR* **28** #122.

[16] MOORE, J. C., Algèbre homologique et homologie espace classificants, *Sem. H. Cartan* 1959/60, Exp. 7.
[17] SERRE, J.-P., Groupes d'homotopie et classes de groupes abéliens, *Ann. of Math.* (2) **58** (1953), 258—294. *MR* **15**—548.
[18] TOLPYGO, A. K., Cohomology of nilpotent Lie algebras and their generating functions, *Uspehi Mat. Nauk* **34** (1979), no. 1, 245—246. *MR* **81d:** 17007.

(Received December 15, 1986)

INSTITUTE OF SOLID STATE PHYSICS OF THE
SOVIET ACADEMY OF SCIENCES
SU—142432 CHERNOGOLOWKA (MOSCOW DISTRICT)
SOVIET UNION

VILLÁNYI ÚT 103
H—1118 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 11—31

MIQUELSCHE MINKOWSKI-EBENEN IN SPIEGELUNGSGEOMETRISCHER DARSTELLUNG

SERGIO CASTRO

Die vorliegende Arbeit enthält eine Charakterisierung der von den Zykelspiegelungen einer miquelschen Minkowski-Ebene erzeugten Automorphismengruppe. Grundlage hierfür ist der Satz von Dienst [3], der diejenigen Minkowski-Ebenen, in denen der Schließungssatz von Miquel gilt, durch die beiden folgenden Eigenschaften kennzeichnet:

(a) An jedem Zykel der Minkowski-Ebene gibt es eine Spiegelung.

(b) Das Produkt dreier Spiegelungen an Zykeln durch zwei nicht parallele Punkte ist eine Zykelspiegelung.

Nachdem Lang [6] durch Spiegelungsrelationen eine Kennzeichnung miquelscher Minkowski-Ebenen einer Charakteristik $\neq 2$ (Charakteristik des der Geometrie zugrundeliegenden Körpers) im Bachmannschen Sinne [1] vornehmen konnte, ist es uns durch einen neuen Ansatz (Grundannahme über eine erzeugte Gruppe) und durch neuartige gruppentheoretische Axiome gelungen, einen spiegelungsgeometrischen Aufbau der genannten Ebenen unabhängig von der Charakteristik anzugeben. Die für miquelsche Möbius-Ebenen durch Wernicke [10] entwickelte spiegelungsgeometrische Darstellung bildete für uns die methodologische Grundlage. Das nun vorliegende Ergebnis ist auch in Einheit mit [10] zu sehen: Für miquelsche Möbius- und Minkowski-Ebenen liegen vergleichbare spiegelungsgeometrische Begründungen vor.

Neben den schon genannten Arbeiten bildeten die Monographie [2] von Benz und die Ergebnisse von Molnár [8], Lang [5], Wernicke [11], sowie von Mäurer [7] und Karzel/Mäurer [4] die wesentlichsten Quellen für unsere Untersuchungen.

Die Anregung zu dieser Arbeit sowie wertvolle Ratschläge für ihre Durchführung verdanke ich Herrn B. Wernicke.

1. Miquelsche Minkowski-Ebenen

1.1. *Minkowski-Ebenen*

Eine *Minkowski-Ebene* (Dienst [3] S. 197) ist eine Inzidenzstruktur $(\mathfrak{P}, \mathfrak{Z}, \in)$, bestehend aus einer Menge $\mathfrak{P}$, deren Elemente *Punkte* heißen, mit zwei *Äquivalenzrelationen* $//_+$ und $//_-$ auf $\mathfrak{P}$ und einer Menge $\mathfrak{Z}$ von Teilmengen von $\mathfrak{P}$, deren Elemente *Zykel* heißen, die folgenden Bedingungen genüge:

1980 *Mathematics Subject Classification* (1985 Revision).Primary 51B10; Secondary 51N15, 51F15.

Key words and phrases. Minkowski planes, cycle reflections in Minkowski planes, characterization of the automorphism group of Minkowski planes with cycle reflections.

M1. Durch drei paarweise nicht parallele Punkte geht genau ein Zykel (Punkte A, B heißen *nicht parallel,* wenn weder $A/\!/_+B$ noch $A/\!/_-B$ gilt).

M2. Zu jedem Punkt A und jedem Zykel z gibt es eindeutig bestimmte Punkte $B, C \in z$ mit $A/\!/_+B$ und $A/\!/_-C$.

M3. Jede Äquivalenzklasse von $/\!/_+$ trifft jede Äquivalenzklasse von $/\!/_-$ in genau einem Punkt.

M4. Es gibt drei paarweise nicht parallele Punkte.

B. (Berühraxiom) Zu jedem Zykel z, jedem Punkt $A \in z$ und jedem zu A nicht parallelen Punkt B gibt es genau einen Zykel z' durch A und B, der z berührt (d. h. $z = z'$ oder $z \cap z' = \{A\}$).

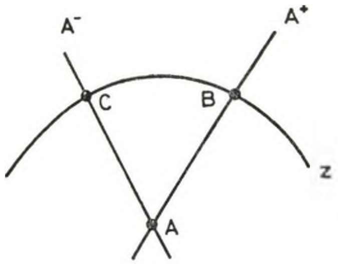

Abb. 1

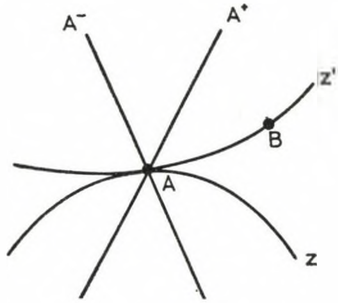

Abb. 2

DEFINITION 1. Für jeden Punkt A einer Minkowski-Ebene sei

$$A^+ := \{X \in \mathfrak{P} : X/\!/_+A\} \quad \text{und} \quad A^- := \{X \in \mathfrak{P} : X/\!/_-A\}.$$

A^+ und A^- heißen *Erzeugende durch A.* Wir bezeichnen mit $\bar{A} = A^+ \cup A^-$ die Menge aller Punkte, die zu A parallel sind. Die Menge aller Äquivalenzklassen bezüglich der Äquivalenzrelation $/\!/_+$ (bzw. $/\!/_-$) wird mit $\mathfrak{E}^+$ (bzw. $\mathfrak{E}^-$) bezeichnet, und $\mathfrak{E}$ ist die Menge aller Erzeugenden. $\blacksquare$

In einer Minkowski-Ebene gelten

(1) *Ist in* M2 *zusätzlich* $A \notin z$, *dann gilt* $B \neq C$. *Wenn* $A \in z$, *dann gilt* $A = B = C$. $\blacksquare$

(2) *Jeder Zykel enthält mindestens drei Punkte.* $\blacksquare$

(3) *Jede Erzeugende enthält mindestens drei Punkte.* $\blacksquare$

(4) *Durch jeden Punkt A geht mindestens ein Zykel.* $\blacksquare$

(5) *Durch je zwei nicht parallele Punkte geht mindestens ein Zykel.* $\blacksquare$

1.2. *Die abgeleiteten Strukturen*

Für jeden Punkt A einer Minkowski-Ebene heißt

$$\mathfrak{A}_A := (\mathfrak{P} \setminus \bar{A}, \{z \setminus \{A\} : A \in z \in \mathfrak{Z}\}, \in)$$

die *im Punkt A abgeleitete Struktur der Minkowski-Ebene.*

Wir betrachten die Menge

$$\mathfrak{G}_A := \{z \setminus \{A\} : A \in z \in \mathfrak{Z}\} \cup \{g \setminus \bar{A} : g \in \mathfrak{E} \setminus \{A^+, A^-\}\}.$$

Die Elemente von $\mathfrak{G}_A$ nennen wir *Geraden.*

Die Struktur

$$\tilde{\mathfrak{A}}_A := (\mathfrak{P} \setminus \bar{A};\ \mathfrak{G}_A;\ \in)$$

heißt *der affine Abschluß der abgeleiteten Struktur* (vgl. Dienst [3] S. 200).

(6) $\mathfrak{A}_A$ *ist eine affine Ebene.* ∎

1.3. *Zykelspiegelungen*

Definition 2. Wir verstehen unter einer *Zykelspiegelung an* $z \in \mathfrak{Z}$ einen von der Identität verschiedenen Automorphismus der Minkowski-Ebene, der z punktweise festläßt (Dienst [3] S. 198). ∎

Eine Zykelspiegelung an z wird mit σ_z bzw. kurz mit σ bezeichnet. Die folgenden Eigenschaften über Zykelspiegelungen lassen sich schrittweise bestätigen:

(7) *Es sei* $A^+ \in \mathfrak{E}^+ (A^- \in \mathfrak{E}^-)$ *und* σ *eine Zykelspiegelung an* $z \in \mathfrak{Z}$. *Dann gilt* $\sigma(A^+) \in \mathfrak{E}$ $(\sigma(A^-) \in \mathfrak{E})$. ∎

(8) *Es sei* σ *eine Zykelspiegelung an* $z \in \mathfrak{Z}$. *Die Fixpunktmenge von* σ *ist* z. ∎

(9) *Ist* σ *eine Zykelspiegelung an* $z \in \mathfrak{Z}$, *dann gilt für alle* $A \in z$, $\sigma(\bar{A}) = \bar{A}$. ∎

(10) *Ist* σ *eine Zykelspiegelung an* $z \in \mathfrak{Z}$ *und* $A \in z$, *dann ist entweder* $\sigma(A^+) = A^+$ *oder* $\sigma(A^+) = A^-$ *(bzw.* $\sigma(A^-) = A^-$ *oder* $\sigma(A^-) = A^+$*).* ∎

(11) *Ist* σ *eine Zykelspiegelung an* $z \in \mathfrak{Z}$ *und* $A \in z$, *dann gilt* $\sigma(A^+) = A^-$ *und* $\sigma(A^-) = A^+$. ∎

Wir können jetzt die Punkt-Bildpunkt-Zuordnung bei einer Zykelspiegelung beschreiben: Ist σ eine Spiegelung an z und $A \notin z$, dann ist $\{\sigma(A)\} = B^- \cap C^+$ mit $\{B\} = A^+ \cap z$ und $\{C\} = A^- \cap z$. Ist $A \in z$, dann ist $\sigma(A) = A$. $\sigma(A)$ heißt der zu A bezüglich z symmetrische Punkt.

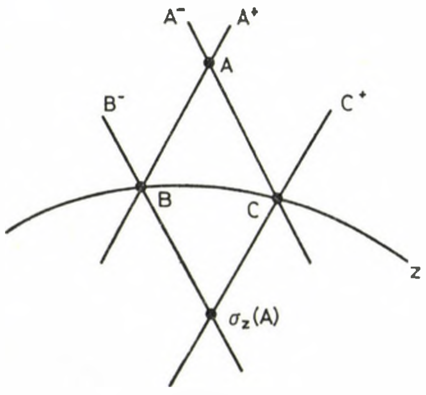

Abb. 3

Neben den Aussagen (1) bis (11) gilt (Dienst [3], S. 200)

(12) *An jedem Zykel z gibt es höchstens eine Spiegelung und diese ist involutorisch. Für jeden Punkt* $A \in z$ *gibt es genau ein Berührbüschel* $\mathfrak{B}$ *in* A, *so daß* $\mathfrak{B} \cup \{z\}$ *die Menge der Fixzykel von* σ_z *durch* A *ist.* ∎

Zwei sich in A berührende Zykel sind in der affinen Ebene $\mathfrak{A}_A$ parallele Geraden. Somit ist das Berühren im Punkt A eine Äquivalenzrelation in der Menge der Zykel durch A. Die Äquivalenzklassen dieser Relation heißen *Berührbüschel in* A.

In weiteren sei $(\mathfrak{P}, \mathfrak{Z})$ stets eine Minkowski-Ebene mit

(i) An jedem Zykel z gibt es eine Spiegelung.

Minkowski-Ebenen mit der Eigenschaft (i) werden wir als *S-Minkowski-Ebenen* bezeichnen.

(13) *Es seien* z, z' *zwei beliebige Zykel. Dann gilt die Relation*

$$\sigma_z \sigma_{z'} \sigma_z = \sigma_{\sigma_z(z')}.$$

Beweis. Es sei $P \in \sigma_z(z')$. Daraus folgt $\sigma_z(P) \in z'$ bzw. $\sigma_z\sigma_{z'}\sigma_z(P) = \sigma_z\sigma_z(P) = P$ nach (8) und (12), d. h., die Punkte auf $\sigma_z(z')$ bleiben bei $\sigma_z\sigma_{z'}\sigma_z$ fest. Es gilt $\sigma_z\sigma_{z'}\sigma_z \neq \mathrm{id}$. In der Tat, wäre $\sigma_z\sigma_{z'}\sigma_z = \mathrm{id}$ dann hätten wir $\sigma_z\sigma_{z'} = \sigma_z$, was nicht geht. Nach (12) ist damit $\sigma_z\sigma_{z'}\sigma_z = \sigma_{\sigma_z(z')}$. ∎

Definition 3. Für Zykel z, z' schreiben wir $z \perp z'$ und sagen z *ist orthogonal zu* z' genau dann, wenn $\sigma_z(z') = z'$ und $z \neq z'$ ist. ∎

Aus (13) folgt unmittelbar für Zykel z, z'

(14) $\sigma_z(z') = z' \Leftrightarrow \sigma_z\sigma_{z'} = \sigma_{z'}\sigma_z$. ∎

(15) *Die folgenden Aussagen sind in S-Minkowski-Ebenen äquivalent:*
- (ii) *Ist* z *ein Zykel und sind* P, Q *Punkte, so daß* $P, Q, \sigma_z(P)$ *paarweise nicht parallel sind, dann liegen* $P, Q, \sigma_z(P), \sigma_z(Q)$ *auf einem Zykel.*
- (iii) *Ist* z *ein Zykel und* P *ein Punkt mit* $P \notin z$, *so gilt für jeden Zykel* z' *mit* $P, \sigma_z(P) \in z'$ *stets* $z' \perp z$.
- (iv) *Ist* z *ein Zykel und sind* P, Q *Punkte, so daß* $P, Q, \sigma_z(P)$ *paarweise nicht parallel*[1] *sind, dann gibt es einen Zykel* z' *mit* $P, Q \in z'$ *und* $z' \perp z$.

Beweis. (ii)⇒(iii). Es gelte (ii), und zu einem Zykel z und zu einem Punkt $P \notin z$ sei z' ein Zykel mit $P, \sigma_z(P) \in z'$. Es sei Q ein Punkt mit $Q \in z'$, $Q \neq P, \sigma_z(P)$. Die Punkte $P, Q, \sigma_z(P), \sigma_z(Q)$ liegen nach (ii) auf einem Zykel, der durch $P, Q, \sigma_z(P)$ eindeutig bestimmt ist. Folglich ist $\sigma_z(z') = z'$.

(iii)⇒(iv). Es gelte (iii). Für P, Q und z seien P, Q und $\sigma_z(P)$ paarweise nicht parallel. Folglich gilt $P \neq \sigma_z(P)$, d. h. $P \notin z$. Wegen (iii) gilt für den Zykel z' durch $P, Q, \sigma_z(P)$ stets $z' \perp z$. Die Rückgewinnung von (ii) aus (iv) ist evident. ∎

---

[1] Da P und $\sigma_z(P)$ nicht parallel sind, ist $\sigma_z(P) \neq P$, d. h. $P \notin z$.

Es sei z ein Zykel, $\mathfrak{B}$ das Berührbüschel in $P \in z$, dem z angehört, und $\mathfrak{B}^*$ das Berührbüschel der Fixzykel von σ_z durch P. Sei $\mathfrak{L}_P$ die Menge der Berührbüschel in P.

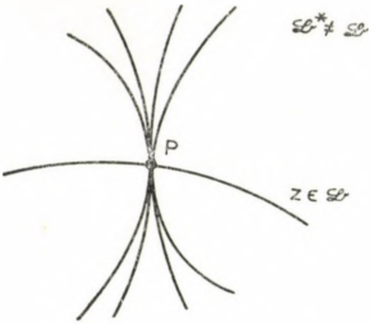

Abb. 4

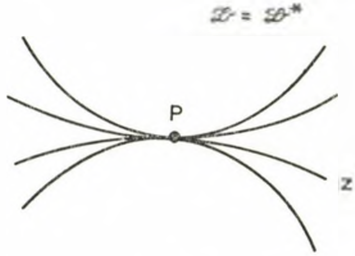

Abb. 5

Wir betrachten die Abbildung

$$\omega_P : \mathfrak{L}_P \to \mathfrak{L}_P \quad \text{mit} \quad \omega_P(\mathfrak{B}) = \mathfrak{B}^*,$$

wobei ω_P^2 nach (14) und (12) die identische Abbildung ist. Zunächst ist in [3] gezeigt worden:

(16) *Die Abbildungen ω_P sind für alle $P \in \mathfrak{P}$ entweder frei von Fixelemente oder gleich der identischen Abbildung auf $\mathfrak{L}_P$.* ∎

Die Ergebnisse von Dienst [3] sind

(17) *Es sei $(\mathfrak{P}, \mathfrak{Z})$ eine S-Minkowski-Ebene, in der $\omega_P \neq \mathrm{id}$ für alle $P \in \mathfrak{P}$ gilt. Dann gilt für Berührbüschel $\mathfrak{B}$, $\mathfrak{B}^*$ in einem Punkt P mit $\sigma_z \sigma_{z^*} = \sigma_{z^*} \sigma_z$ für $z \in \mathfrak{B}$, $z^* \in \mathfrak{B}^*$ stets $\mathfrak{B} \neq \mathfrak{B}^*$. Zu jedem Punkt $A \in \mathfrak{P} \setminus \bar{P}$ gibt es Zykel $y \in \mathfrak{B}$ und $y^* \in \mathfrak{B}^*$ durch A derart, daß $\sigma_y \sigma_{y^*}|_{\mathfrak{P} \setminus P}$ eine involutorische Streckung (Punktspiegelung) von $\bar{\mathfrak{A}}_P$ mit A als Zentrum ist. $\bar{\mathfrak{A}}_P$ ist eine Translationsebene einer Charakteristik $\neq 2$.* ∎

(18) *Es sei $(\mathfrak{P}, \mathfrak{Z})$ eine S-Minkowski-Ebene, in der $\mathfrak{B} = \mathfrak{B}^*$. Seien $y, y^* \in \mathfrak{B}$. Dann gilt entweder $\sigma_y \sigma_{y^*}$ ist involutorisch und hat P als einzigen Fixpunkt oder aber $\sigma_y \sigma_{y^*} = \mathrm{id}$, falls $y = y^*$. $\bar{\mathfrak{A}}_P$ ist eine Translationsebene der Charakteristik 2.* ∎

Wir nennen eine *Minkowski-Ebene von Charakteristik* $\neq 2$ *bzw.* 2, je nachdem, ob $\bar{\mathfrak{A}}_P$ für ein $P \in \mathfrak{P}$ eine Translationsebene der Charakteristik $\neq 2$ bzw. 2 ist (vgl. Dienst [3] S. 202).

In S-Minkowski-Ebene einer Charakteristik $\neq 2$ gilt nach (17)

(19) *Zwei orthogonale Zykel durch einen Punkt haben einen weiteren gemeinsamen Punkt, und für zwei verschiedene Punkte eines Zykels gibt es stets einen Zykel durch diese Punkte, der orthogonal zum gegebenen Zykel ist.* ∎

1.4. *Miquelsche Minkowski-Ebenen*

Eine Minkowski-Ebene heißt *miquelsch,* wenn sie zusätzlich dem Satz von Miquel genügt:

SATZ VON MIQUEL. *Es seien* A_1, A_2, A_2, A_4, B_1, B_2, B_3, B_4 *acht paarweise nicht parallele Punkte. Wenn dann die Punktquadrupel* (A_1, A_2, B_1, B_2); (A_2, A_3, B_2, B_3); (A_3, A_4, B_3, B_4); (A_4, A_1, B_4, B_1); (A_1, A_2, A_3, A_4) *jeweils auf einem Zykel liegen, so gilt dies auch für das Punktquadrupel* (B_1, B_2, B_3, B_4).

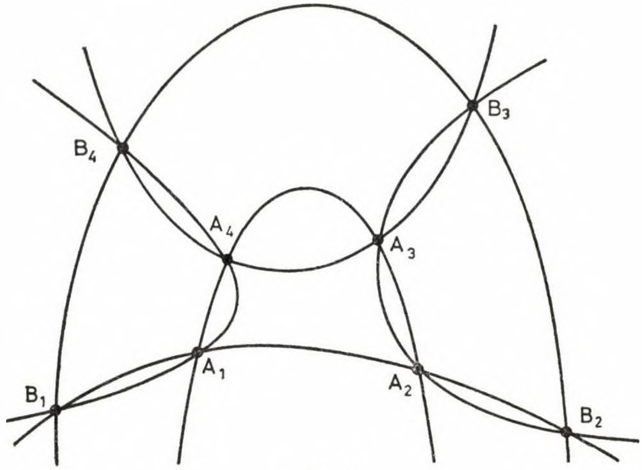

Abb. 6

Die miquelsche Minkowski-Ebenen sind durch folgenden Satz charakterisiert (Dienst [3] S. 199).

SATZ VON DIENST. *Minkowski-Ebenen, die den Eigenschaften*

a) *An jedem Zykel gibt es eine Spiegelung;*

b) *Das Produkt dreier Spiegelungen an Zykeln eines Zykelbüschels (Menge aller Zykel durch zwei feste Punkte) läßt einen Zykel punktweise fix*[2]

genügen, sind bis auf Isomorphie genau die miquelschen Minkowski-Ebenen. ▮

2. Gruppentheoretisches Axiomensystem für miquelsche Minkowski-Ebenen

2.1. *Axiomensystem und erste Folgerungen*

GRUNDANNAHME. Gegeben sei eine Gruppe **G**, die von einer Menge $\mathfrak{Z}=\{\alpha, \beta, \gamma, \ldots\}$ erzeugt wird. Und es gibt ein System $\mathscr{P}=\{A, B, C, \ldots\}$ von Teilmengen von $\mathfrak{Z}$ mit den Eigenschaften:

---

[2] Aus Eigenschaften über Zykelspiegelungen in Minkowski-Ebenen mit a) ergibt sich, daß das Produkt genau die Punkte eines Zykels des Büschels festläßt, d. h., es ist eine Zykelspiegelung (vgl. [3]).

(I) Für alle $A \in \mathscr{P}$ und für alle $\alpha \in \mathfrak{Z}$ gilt

$$\alpha \in A \Leftrightarrow A^\alpha := \alpha^{-1} A \alpha = A.$$

(II) Für alle $\alpha \in \mathfrak{Z}$ und für alle $A \in \mathscr{P}$ ist $A^\alpha \in \mathscr{P}$.
(III) Für jedes $\alpha \in \mathfrak{Z}$ gibt es ein $A \in \mathscr{P}$ mit $\alpha \in A$. ∎

Die Eigenschaften (I), (II) und (III) von $\mathscr{P}$ besagen, daß $\mathscr{P}$ eine aus selbstinvarianten Mengen bestehende, invariante Überdeckung von $\mathfrak{Z}$ ist.

Definition 1. Die Elemente von $\mathscr{P}$ heißen *Punkte*. ∎

Definition 2. In der Menge $\mathfrak{P} \times \mathfrak{Z}$ sei die Relation „I“ bestimmt durch

$$A \, \mathrm{I} \, \alpha :\Leftrightarrow \alpha \in A. \; ∎$$

Definition 3. In der Menge $\mathbf{G} \times \mathbf{G}$ sei die Relation „|“ bestimmt durch

$$a|b :\Leftrightarrow a, b, ab \text{ involutorisch.} \; ∎$$

Bemerkung 1. In der Menge der involutorischen Elemente aus $\mathbf{G}$ ist $a|b$ äquivalent mit

$$(ab)^2 = 1 \quad \text{und} \quad ab \neq 1$$

bzw. mit

$$ab = ba \quad \text{und} \quad a \neq b$$

bzw. mit

$$bab =: a^b = a \quad \text{und} \quad a \neq b.$$

Die |-Relation ist symmetrisch, und es folgt insbesondere aus $a|b$ stets $a \neq b$.

Definition 4. Punkte A, B heißen genau dann *verbindbar*, wenn $A \cap B \neq \emptyset$ gilt. Ist $A \cap B = \emptyset$ so heißen A und B *unverbindbar*. Punkte A, B heißen genau dann *parallel* (in Zeichen $A//B$), wenn $A = B$ oder $A \cap B = \emptyset$. ∎

Der erzeugten Gruppe $(\mathbf{G}, \mathfrak{Z})$ wird eine geometrische Struktur $(\mathscr{P}, \mathfrak{Z})$ bestehend aus der Menge $\mathscr{P} = \{A, B, C, \ldots\}$ der Punkte und der Menge $\mathfrak{Z}$ der Erzeugenden[3], hierbei Zykel genannt, zugeordnet. $(\mathscr{P}, \mathfrak{Z})$ heißt die *Gruppenebene von* $(\mathbf{G}, \mathfrak{Z})$.

Die erzeugte Gruppe genüge neben der Grundannahme den folgenden *Axiomen:*

A1. Es gibt drei[4] paarweise verbindbare Punkte.

A2. Zu drei paarweise verbindbaren Punkten A, B, C gibt es genau eine Erzeugende α mit $A, B, C \, \mathrm{I} \, \alpha$.

A3. Zu einer Erzeugenden α und zu einem Punkt A mit $A \mathrm{I} \alpha$ gibt es einen Punkt $B \not\mathrm{I} \alpha$, der mit A verbindbar ist.

A4. Zu zwei verbindbaren Punkten A, B gibt es genau zwei Punkte, die weder mit A noch mit B verbindbar sind.

A5. (Minkowski-Axiom) Zu einem Punkt A und einer Erzeugenden α mit $A \not\mathrm{I} \alpha$ gibt es genau zwei Punkte B, C mit $B, C \, \mathrm{I} \, \alpha$, so daß A, B und A, C unverbindbar sind.

---

[3] Verwechslungen mit dem Begriff „Erzeugende“ in der Inzidenzstruktur $(\mathfrak{P}, \mathfrak{Z})$, vgl. Abschnitt 1, sind wohl nicht zu befürchten.

[4] Die Kardinalzahl schließt im weiteren die Formulierung „paarweise verschieden“ mit ein.

A6. Zu drei Erzeugenden α, β, γ und zwei Punkten A, B mit $A, B\,\mathrm{I}\,\alpha, \beta$ und $A\,\mathrm{I}\,\gamma$ gibt es einen Punkt C mit $C\,\mathrm{I}\,\gamma$, $C \neq A$ derart, daß $C\,\mathrm{I}\,\alpha$ oder $C\,\mathrm{I}\,\beta$ gilt.[5]

A7. Sind A, B zwei Punkte und α, β, γ Erzeugende mit $A, B\,\mathrm{I}\,\alpha, \beta, \gamma$ und $\alpha|\beta, \gamma$, dann folgt $\beta=\gamma$.

A8. Zu einer Erzeugenden α und zu Punkten A, B, so daß $A \not\mathrm{I}\,\alpha$ und B mit A und A^α verbindbar ist, gibt es eine Erzeugende β mit $A, B\,\mathrm{I}\,\beta$ und $\beta|\alpha$.

Bemerkung 2. Da $\mathscr{P}$ nach (III) in der Grundannahme eine Überdeckung von $\mathfrak{Z}$ und $\mathscr{P}$ wegen der Eigenschaft (II) invariant gegenüber inneren Automorphismen von $\mathbf{G}$ ist, folgt somit, daß auch $\mathfrak{Z}$ invariant gegenüber inneren Automorphismen von $\mathbf{G}$ ist.

Bemerkung 3. Wie man nun leicht nachrechnet, ist die I-Relation ebenfalls invariant gegenüber inneren Automorphismen von $\mathbf{G}$: Aus $A\,\mathrm{I}\,\alpha$ folgt $A^\beta\,\mathrm{I}\,\alpha^\beta$ für alle $\beta \in \mathfrak{Z}$.

Für die der erzeugten Gruppe $(\mathbf{G}, \mathfrak{Z})$, die der Grundannahme und den Axiomen A1 bis A8 genügt, zugeordnete Gruppenebene $(\mathscr{P}, \mathfrak{Z})$ ergeben sich wichtige Konsequenzen.

Wir benutzen die Sprache der Gruppenebene und nehmen die Relation „I" zwischen Punkten und Erzeugenden als *Inzidenzrelation* in $\mathscr{P} \times \mathfrak{Z}$. Wir sagen weiter, daß sich zwei Erzeugende α, β aus $\mathfrak{Z}$ *berühren*, wenn es genau einen Punkt A mit $A\,\mathrm{I}\,\alpha, \beta$ gibt.

(1) *Jede Erzeugende α indiziert mit mindestens drei Punkten.*

Beweis. Es sei α eine Erzeugende. Da $\mathscr{P}$ eine Überdeckung von $\mathfrak{Z}$ ist, folgt die Existenz eines Punktes A mit $A\,\mathrm{I}\,\alpha$. Nach A3 gibt es einen Punkt $B \not\mathrm{I}\,\alpha$, der mit A verbindbar ist. Nach A5 gibt es genau zwei Punkte C, D mit $C, D\,\mathrm{I}\,\alpha$, so daß B, C und B, D unverbindbar sind. Es ist $C, D \neq A$, weil A mit B verbindbar ist. Folglich sind A, C, D drei verschiedene Punkte auf α. ▮

Mit (1) erhalten wir eine dem Axiom A3 verwandte Aussage.

(2) *Zu einer Erzeugenden α und zu einem Punkt A mit $A \not\mathrm{I}\,\alpha$ gibt es einen Punkt $B\,\mathrm{I}\,\alpha$, der mit A verbindbar ist.*

Beweis. Er ergibt sich unmittelbar aus A5 und (1), da es zu A nach A5 genau zwei unverbindbare Punkte auf α gibt und α nach (1) mindestens drei Punkte enthält. ▮

(3) *Zu jeder Erzeugenden α und zu jedem Punkt A mit $A\,\mathrm{I}\,\alpha$ gibt es eine Erzeugende β mit $A\,\mathrm{I}\,\beta$ und $\beta|\alpha$.*

Beweis. Es seien α, A mit $A\,\mathrm{I}\,\alpha$ gegeben. Nach A3 gibt es einen Punkt B mit $B \not\mathrm{I}\,\alpha$, und A, B sind verbindbar. Aus der Verbindbarkeit von A und B folgt die von A und B^α wegen $A\,\mathrm{I}\,\alpha$. Für A, B und B^α treffen die Voraussetzungen von A8 zu, so daß es ein β mit $A, B\,\mathrm{I}\,\beta$ und $\beta|\alpha$ gibt. ▮

---

[5] Dieses Axiom wurde in [10] bei einer spiegelungsgeometrischen Darstellung Möbiusscher Ebenen benutzt. Deutet man A6 in der in A abgeleiteten Struktur $\mathfrak{A}_A$ (bzw. in ihrem affinen Abschluß $\tilde{\mathfrak{A}}_A$) einer Minkowski-Ebene, so tritt der grundlegende „affine" Charakter von A6 hervor.

Für jede Erzeugende α gibt es einen Punkt A mit $A\mathrm{I}\alpha$, und somit erhalten wir als Folgerung aus (3) die Ergänzung zur Grundannahme

(4) *Das Erzeugendensystem $\mathfrak{Z}$ besteht aus involutorischen Elementen.* ∎

(5) *Die Punkte A und A^α sind für alle Punkte A und Erzeugenden α stets verbindbar.*

Beweis. Es sei $A\not\mathrm{I}\alpha$. Nach (2) gibt es einen Punkt $B\mathrm{I}\alpha$, der mit A verbindbar ist. Wegen $B\mathrm{I}\alpha$ ist auch B mit A^α verbindbar. Nach A8 gibt es nun ein β mit $A, B\mathrm{I}\beta$ und $\beta|\alpha$. Da $A\mathrm{I}\beta$ und $\beta|\alpha$ gilt, ist $A^\alpha\mathrm{I}\beta^\alpha=\beta$. Folglich ist A mit A^α verbindbar. ∎

Bemerkung 4. Nach (1) inzidiert jede Erzeugende mit mindestens drei Punkten. Anderseits gibt es zu drei paarweise verbindbaren Punkten genau eine Erzeugende, die mit diesen Punkten inzidiert. Sind A, B, C drei paarweise verbindbare Punkte, so bezeichnen wir die Erzeugende α mit $A, B, C\mathrm{I}\alpha$ auch durch (A, B, C): $\alpha=(A, B, C)$.

2.2. *Berührsatz und Parallelitätsrelationen in der Gruppenebene*

Für die Gruppenebene $(\mathscr{P}, \mathfrak{Z})$ können wir das Berühraxiom B aus 1.1 nachweisen und nach Einführung zweier Parallelitätsrelationen in der Menge $\mathscr{P}$ der Punkte zeigen, daß die Gruppenebene eine Minkowski-Ebene ist.

(6) *Sind A, B verbindbare Punkte und ist α eine Erzeugende mit $A\mathrm{I}\alpha$ und $B\not\mathrm{I}\alpha$, so gibt es genau eine Erzeugende β mit $A, B\mathrm{I}\beta$ und α, β berühren sich.*

Beweis. Es seien A, B zwei verbindbare Punkte und α eine Erzeugende mit $A\mathrm{I}\alpha$ und $B\not\mathrm{I}\alpha$ gegeben.

Existenz. Wir beweisen zuerst, daß A, B und B^α drei paarweise verschiedene und verbindbare Punkte sind. Aus $A\mathrm{I}\alpha$ und $B\not\mathrm{I}\alpha$ folgt $A^\alpha=A$, $B^\alpha\neq B$ und $B^\alpha\not\mathrm{I}\alpha$. Daher sind A, B, B^α drei paarweise verschiedene Punkte. Nach (5) sind B und B^α verbindbar. Da A, B als verbindbar vorausgesetzt wurden, existiert eine Erzeugende α' mit $A, B\mathrm{I}\alpha'$. Dann gilt $A=A^\alpha$, $B^\alpha\mathrm{I}(\alpha')^\alpha\in\mathfrak{Z}$, d. h., A und B^α sind verbindbar. Nun existiert nach A2 eine Erzeugende γ mit $A, B, B^\alpha\mathrm{I}\gamma$. Wir haben $\gamma^\alpha=(A, B, B^\alpha)^\alpha=$ $=(A, B^\alpha, B)=\gamma$ und wegen $\gamma\neq\alpha$ folgt $\gamma|\alpha$.

Fall 1. Gibt es außer A, B, B^α keinen weiteren Punkt P mit $P\mathrm{I}\gamma$, so ist mit $\beta:=\gamma$ die Existenzaussage bestätig.

Fall 2. Wir setzen jetzt voraus, daß es einen Punkt $C\neq A, B, B^\alpha$ mit $C\mathrm{I}\gamma$ gibt.

Die Idee der Demonstration ist dann folgende: Wir zeigen zunächst, daß es einen Punkt E mit $E\not\mathrm{I}\alpha, \gamma$ gibt, so daß E und A verbindbar sind. Danach bestimmen wir die Erzeugende $\varepsilon=(A, E, E^\alpha)$ und beweisen, daß die Erzeugende (A, B, B^ε) die Erzeugende α in A berührt.

Aus A3 folgt die Existenz eines Punktes D mit $D\not\mathrm{I}\gamma$ und A, D verbindbar. Ist $D\not\mathrm{I}\alpha$, so sei $E:=D$.

In weiteren sei $D\mathrm{I}\alpha$. Nach A5 ist dann D mindestens mit einem der Punkte C, B, B^α verbindbar.

a) D sei mit B (und folglich mit B^α) verbindbar.

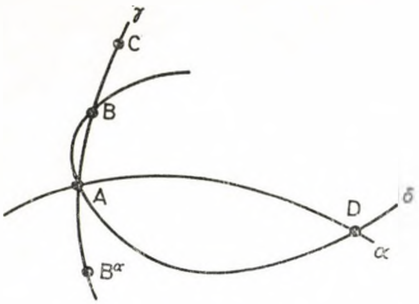

Abb. 7

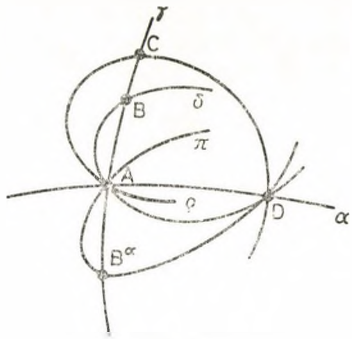

Abb. 8

Wir betrachten den Zykel δ mit $A, B, D \mathrm{I} \delta$. Offenbar gilt $B^{\alpha}, C \not\mathrm{I} \delta$. Wir bewiesen jetzt, daß für die Punkte $B^{\alpha\delta}$ und C^{δ} gilt:
1) $B^{\alpha\delta}$ und C^{δ} sind mit A verbindbar.
2) $B^{\alpha\delta}, C^{\delta} \not\mathrm{I} \gamma$.
3) Es gilt nicht gleichzeitig $B^{\alpha\delta} \mathrm{I} \alpha$ und $C^{\delta} \mathrm{I} \alpha$.

Zu 1). Da $\gamma = (C, A, B^{\alpha})$ und $A \mathrm{I} \delta$ ist, gilt $\gamma^{\delta} = (C^{\delta}, A, B^{\alpha\delta})$.

Zu 2). Angenommen, es gilt $B^{\alpha\delta} \mathrm{I} \gamma$. Es würde $\gamma^{\delta} = (A, B, B^{\alpha})^{\delta} = (A, B, B^{\alpha\delta}) = \gamma$ folgen, und mit $B^{\alpha} \not\mathrm{I} \delta$ hätten wir noch $\gamma \neq \delta$. Es gilt danach $\gamma | \delta$ bzw. auch $\delta^{\gamma} = \delta$. Da $\gamma^{\alpha} = (A, B, B^{\alpha})^{\alpha} = (A, B^{\alpha}, B) = \gamma$ ist, gilt $\alpha^{\gamma} = \alpha$. Aus $D \mathrm{I} \alpha, \delta$ folgte $D^{\gamma} \mathrm{I} \alpha, \delta$. Neben A und D wäre D^{γ} ein dritter Punkt von α und δ, was $\alpha = \delta$ zur Folge hätte. Genauso beweist man $C^{\delta} \not\mathrm{I} \gamma$.

Zu 3). Seien $B^{\alpha\delta}, C^{\delta} \mathrm{I} \alpha$. Da A, B^{α}, D drei verschiedene, paarweise verbindbare Punkte sind, sei $\pi := (A, B^{\alpha}, D)$. Es gilt $C \not\mathrm{I} \delta$. Dann sind A, D, C^{δ} drei verschiedene Punkte. Nach Voraussetzung ist $C^{\delta} \mathrm{I} \alpha$ und folglich ist C^{δ} mit A und D verbindbar.

Sei $(A, D, C^{\delta})^{\delta} = (A, D, C) = \varrho$. Wir haben $\pi^{\delta} = (A, B^{\alpha}, D)^{\delta} = (A, B^{\alpha\delta}, D) = \alpha$; $\varrho^{\delta} = (A, D, C)^{\delta} = (A, D, C^{\delta}) = \alpha$, und folglich $\varrho = \alpha^{\delta} = \pi$, was nicht sein kann. Also gilt $B^{\alpha\delta} \not\mathrm{I} \gamma$ oder $C^{\delta} \not\mathrm{I} \gamma$, womit ein Punkt gefunden wäre, der weder mit α noch mit γ inzidiert und mit A verbindbar ist.

b) D sei mit B nicht verbindbar. Dann ist auch D mit B^{α} nicht verbindbar, und nach A5 ist D mit C verbindbar.

b1) Es sei $C \not\mathrm{I} \alpha$. Für die Erzeugende $\delta = (A, C, D)$ und für B gilt $B \not\mathrm{I} \delta$, und für B^{δ} erhalten wir schließlich:
1) B^{δ} ist mit A verbindbar.
2) $B^{\delta} \not\mathrm{I} \alpha, \gamma$.

Zu 1). Es gilt $B, A \mathrm{I} \gamma$ und $A \mathrm{I} \delta$. Daraus folgt $B^{\delta}, A \mathrm{I} \gamma^{\delta}$, d. h., B^{δ} ist mit A verbindbar.

Zu 2). Wir setzen $B^{\delta} \mathrm{I} \gamma$ voraus. Aus $B \not\mathrm{I} \delta$ folgt $B^{\delta} \not\mathrm{I} \delta$, und folglich ist $B^{\delta} \neq C, A$. Es gilt $\gamma^{\delta} = (A, C, B^{\delta})^{\delta} = (A, C, B) = \gamma$ und $\gamma \neq \delta$. Daraus folgt $\gamma | \delta$,

d. h. $\delta^\gamma = \delta$. Wegen $\gamma^\alpha = (A, B, B^\alpha)^\alpha = \gamma$ ist $\alpha^\gamma = \alpha$. Mit $D\mathrm{I}\alpha, \delta$ gilt nun auch $D^\gamma \mathrm{I}\alpha, \delta$. Aus $A, D, D^\gamma \mathrm{I}\alpha, \delta$ folgt $\alpha = \delta$, was nicht geht wegen $C\not{\mathrm{I}}\alpha$ und $C\mathrm{I}\delta$. Also gilt $B^\delta \not{\mathrm{I}}\gamma$.

Wir setzen $B^\delta \mathrm{I}\alpha$ voraus. Dann gilt $B^\delta \neq A, D$ wegen $B\not{\mathrm{I}}\delta$. Folglich ist $\alpha^\delta = (A, B^\delta, D)^\delta = (A, B, D) \in \mathfrak{Z}$. Das steht im Widerspruch dazu, daß B und D unverbindbar sind. Es gilt auch $B^\delta \not{\mathrm{I}}\alpha$.

In diesem Fall sei $E := B^\delta$.

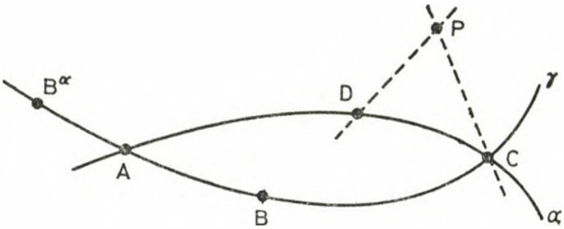

Abb. 9

b2) Es sei $C\mathrm{I}\alpha$. Nach A4 gibt es einen Punkt P, der weder mit C noch mit D verbindbar ist. P liegt weder auf α noch auf γ, weil sonst C und P verbindbar wären. Nach A5 ist P mit A verbindbar. Hierbei erfüllt $E := P$ die Forderungen bezüglich α und γ.

Wir haben jetzt neben α, A, B eine Erzeugende γ mit $A, B\mathrm{I}\gamma$ und $\gamma|\alpha$ und einen Punkt E mit $E\not{\mathrm{I}}\alpha, \gamma$, so daß E mit A verbindbar ist. Nach (5) und A2 bestimmen A, E, E^α eine Erzeugende ε, und wegen $\varepsilon^\alpha = (A, E, E^\alpha)^\alpha = (A, E^\alpha, E) = \varepsilon$ und $\varepsilon \neq \alpha$ gilt $\varepsilon|\alpha$.

Es ist $B\not{\mathrm{I}}\varepsilon$. Wäre $B\mathrm{I}\varepsilon$, so folgte aus $\varepsilon|\alpha$ stets $B^\alpha \mathrm{I}\varepsilon$ bzw. $\varepsilon = (A, E, E^\alpha) = (A, B, B^\alpha) = \gamma$, was wegen $E\not{\mathrm{I}}\gamma$ nicht geht.

Sei $\beta = (A, B, B^\varepsilon)$. β berührt α in Punkt A. In der Tat, setzen wir nämlich voraus, daß ein Punkt F existiert mit $F \neq A$ und $F\mathrm{I}\alpha, \beta$, so erhalten wir einen Widerspruch. Falls $F\not{\mathrm{I}}\varepsilon$ ist, folgt wegen $\alpha, \beta|\varepsilon$, daß A, F und F^ε sowohl auf α als auch auf β liegen. Das hätte $\alpha = \beta$ zur Folge, im Widerspruch zu $B\not{\mathrm{I}}\alpha$ und $B\mathrm{I}\beta$.

Ist $F\mathrm{I}\varepsilon$, dann folgt aus $A, F\mathrm{I}\varepsilon, \alpha, \beta$ und $\varepsilon|\alpha, \beta$ nach A7 ebenfalls $\alpha = \beta$. Insgesamt hat die durch A, B, B^ε bestimmte Erzeugende β mit α genau den Punkt A gemeinsam.

E i n d e u t i g k e i t. Neben einer Erzeugenden β, die α im A berührt und durch $B\not{\mathrm{I}}\alpha$ geht, betrachten wir eine Erzeugende β' mit $A, B\mathrm{I}\beta'$ und $\beta' \neq \beta$. Für β' gilt nun zusammen mit den übrigen Punkten und Erzeugenden $A, B\mathrm{I}\beta, \beta'$ und $A\mathrm{I}\alpha$, wobei α, β, β' paarweise verschiedenen sind und $A \neq B$ ist. Da sich weiterhin α und β berühren, gibt es nach A6 einen Punkt C mit $C\mathrm{I}\alpha, \beta'$ und $C \neq A$. Folglich ist β die einzige Erzeugende durch B, die α im A berührt. ▮

Die folgenden Sätze (7) und (8) sind Hilfssätze über verbindbare bzw. unverbindbare Punkte, die wir zur Ableitung von Eigenschaften über isotrope Geraden, vgl. die Definition 5 vor (9), benötigen.

(7) *Sind A, B, C drei paarweise verbindbare Punkte und α die Erzeugende durch A, B, C, dann gibt es genau eine Erzeugende β mit $B\mathrm{I}\beta$, $\beta|\alpha$ und $A^{\beta}=C$.*

Beweis. Es seien für A, B, C und α die Voraussetzungen des Satzes erfüllt. Nach A4 gibt es einen Punkt P, der weder mit A noch mit C verbindbar ist. Nach A5 ist P mit B verbindbar. B, P, P^{α} sind drei paarweise verbindbare Punkte. Sei $\beta=(P, P^{\alpha}, B)$. Es ist $\beta\neq\alpha$, und es gilt $\beta^{\alpha}=(P, P^{\alpha}, B)^{\alpha}=(P^{\alpha}, P, B)=\beta$, d. h., $\beta|\alpha$. Da C mit P unverbindbar ist, gilt $C\not\mathrm{I}\beta$ und folglich $C^{\beta}\neq C$. C ist nicht mit P^{α} verbindbar, da sonst auch C und P verbindbar wären. Da $\beta|\alpha$ und $P, P^{\alpha}\mathrm{I}\beta$ ist, gilt $C^{\beta}\mathrm{I}\alpha$ und auch die Unverbindbarkeit von C^{β}, P und C^{β}, P^{α}. Wegen $C^{\beta}\neq C$ folgt $C^{\beta}=A$ nach A4.

Ist δ eine von β verschiedene Erzeugende mit $B\mathrm{I}\delta$, $\delta|\alpha$ und $A^{\delta}=C$, dann folgt zunächst $A, C\not\mathrm{I}\delta$. Zu A, δ gibt es nach A5 genau zwei Punkte $D, E\mathrm{I}\delta$, die beide mit A unverbindbar sind. Da $\delta\neq\beta$ ist, gilt $\{D, E\}\neq\{P, P^{\alpha}\}$. Wegen $\delta|\alpha$ und $A^{\delta}=C$ sind auch C, D bzw. C, E unverbindbar. Zu A, C haben wir also mehr als zwei Punkte aufgezeigt, die sowohl mit A als auch mit C unverbindbar sind. Dies widerspricht dem Axiom A4. ∎

(8) *Sind A, B, C drei Punkte, wobei A und B verbindbar sind und C mit A und mit B unverbindbar ist, und ist γ eine Erzeugende mit $A^{\gamma}=B$, so folgt $C\mathrm{I}\gamma$.*

Beweis. Es seien A, B verbindbar und C weder mit A noch mit B verbindbar. Sei γ eine Erzeugende mit $A^{\gamma}=B$. Nach A5 gibt es genau zwei Punkte D und E auf γ, die mit A unverbindbar sind. Dann sind auch D und E mit B unverbindbar. Wäre D mit B verbindbar und δ eine Erzeugende durch D, B, dann folgte $D, A\mathrm{I}\delta^{\gamma}$ wegen $D^{\gamma}=D$, $B^{\gamma}=A$ und $D, B\mathrm{I}\delta$, was nicht sein kann. So sind D und E die zwei nach A4 bestimmten Punkte mit $D, E//A, B$. Daraus folgt $C=D$ oder $C=E$, d. h. $C\mathrm{I}\gamma$. ∎

Definition 5. Sind A, B zwei unverbindbare Punkte, so heißt

$$i(A, B):=\{X\colon A, B//X\}$$

die *durch A, B bestimmte isotrope Gerade.* ∎

(9) *Aus $A//B$, $A\neq B$ und $C, D\in i(A, B)$ folgt $C//D$ sowie $i(A, B)=i(C, D)$, falls $C\neq D$.*

Beweis. Es gelte $C, D\in i(A, B)$ und $C\neq D$. Angenommen es wäre $C \not\!\!/\!\!/ D$, d. h., C, D sind verbindbar. Sei α eine Erzeugende durch C, D. Nach (1) enthält α einen dritten Punkt P. Wegen (7) existiert eine Erzeugende β mit $P\mathrm{I}\beta$ und $C^{\beta}=D$. Aus (8) folgt $A, B\mathrm{I}\beta$ im Widerspruch zu $A//B$ und $A\neq B$. Daraus folgert man $i(A, B)=i(C, D)$: In der Tat, zunächst gilt $A, B\in i(C, D)$. Es sei nun $X\in i(C, D)$ und $X\neq A, B$. Wäre $X \not\!\!/\!\!/ B$, dann existiert eine Erzeugende γ durch B und X. γ enthält nach (1) einen dritten Punkt K. Nach (7) existiert δ mit $K\mathrm{I}\delta$, so daß $B^{\delta}=X$. Aus (8) folgt $C, D\mathrm{I}\delta$, was nicht geht wegen $C\neq D$ und $C//D$. Genauso beweist man $X//A$. Folglich ist $X\in i(A, B)$. Gleichfalls erhält man: Aus $X\in i(A, B)$ mit $A\neq B$ folgt $X\in i(C, D)$. ∎

(10) *Durch jeden Punkt gehen genau zwei verschiedene isotrope Geraden.*

Beweis. Es sei P ein beliebiger Punkt. Nach A1 gibt es drei paarweise verbindbare Punkte P_1, P_2, P_3. Sei α die Erzeugende durch diese Punkte.

a) $P \not{\mathrm{I}}\, \alpha$. Nach A5 gibt es auf α genau zwei Punkte A, B, die mit P unverbindbar sind. Sei $Q /\!/ P$ und $Q \neq A, B, P$. Wegen A5 ist $Q \not{\mathrm{I}}\, \alpha$. Es gilt auch $Q \neq P^\alpha$, weil P mit P^α verbindbar ist. Nach A4 ist Q nicht zugleich parallel zu A und B (P und P^α sind genau die zwei Punkte, die mit A und B unverbindbar sind).

Ist $Q /\!/ A$, dann folgt $Q \in i(P, A)$ wegen $P /\!/ Q$. Wäre Q weder zu A noch zu B parallel, dann wäre P parallel zu drei Punkten auf der Erzeugenden durch A, B, Q, was A5 widerspricht.

b) $P\,\mathrm{I}\,\alpha$. Sei o. B. d. A. $P \neq P_1$. Nach A4 existieren Punkte A, B mit $A, B /\!/ P, P_1$. Wegen der Eindeutigkeitsaussage in A4 gilt $A^\alpha = B$, und folglich gibt es nach (5) eine Erzeugende β durch A und B. Nun sei $Q /\!/ P$ und $Q \neq A, B, P$. Wegen A5 ist $Q \not{\mathrm{I}}\, \beta$. Es gilt $Q \neq P^\beta$, weil P^β mit P verbindbar ist. Nach A4 ist Q nicht zugleich parallel zu A und B. Ist $Q /\!/ B$, dann folgt $Q \in i(P, B)$ wegen $P /\!/ Q$. Ist $Q /\!/ A$, dann folgt $Q \in i(P, A)$ wegen $Q /\!/ P$. Wäre Q weder zu A noch zu B parallel, dann wäre P parallel zu drei Punkten auf der Erzeugenden durch A, B, Q, was A5 widerspricht. ∎

Definition 6. R sei die Menge aller isotropen Geraden, und i_1, i_2 seien die beiden isotropen Geraden durch einen Punkt P. Ferner setzen wir fest:

$$R_1 := \{r \in R : |r \cap i_2| = 1\}; \quad R_2 := \{r \in R : |r \cap i_1| = 1\}.^{6} \quad ∎$$

(11) a) *Sind* r_k, r'_k *zwei verschiedene isotrope Geraden aus* R_k, *so ist* $|r_k \cap r'_k| = 0$, $(k = 1, 2)$.

b) *Ist* $r_1 \in R_1$, $r_2 \in R_2$, *so folgt* $|r_1 \cap r_2| = 1$.

c) *Es ist* $R = R_1 \cup R_2$ *und* $R_1 \cap R_2 = \emptyset$.

Beweis. Zu a). Seien r_1, r'_1 zwei verschiedene isotrope Geraden aus R_1. Seien $\{P_1\} = r_1 \cap i_2$, $\{P'_1\} = r'_1 \cap i_2$. Es gilt $P_1 \neq P'_1$. Wäre $P_1 = P'_1$, dann wären r_1, r'_1, i_2 drei verschiedene isotrope Geraden durch $P_1 = P'_1$, was (10) widerspricht. Gäbe es einen Punkt $A \in r_1, r'_1$, so ist $A \neq P_1, P'_1$. Es folgte aber $A /\!/ P_1, P'_1$ und damit $A \in i(P_1, P'_1) = i_2$ in Widerspruch zu $A \in r_1$, $A \neq P_1$ und $r_1 \cap i_2 = \{P_1\}$.

Zu b). Seien $r_1 \in R_1$, $r_2 \in R_2$, $\{A\} = r_1 \cap i_2$, $\{B\} = r_2 \cap i_1$. O. B. d. A. sei $A, B \neq P \in i_1, i_2$. Es gilt $A \nparallel B$: Wegen $P /\!/ B$ folgte ansonsten mit $B \in i(P, A)$ bzw. $i_1 = i_2$ ein Widerspruch. Sei α mit $A, B\,\mathrm{I}\,\alpha$. Wegen $P /\!/ A, B$ folgt $P^\alpha /\!/ A, B$. Es ist $P^\alpha \notin i_1, i_2$, da sonst $P /\!/ P^\alpha$ wäre. Wir haben $i_1^\alpha = (P, B)^\alpha = (P^\alpha, B) = r_2$, $i_2^\alpha = (P, A)^\alpha = (P^\alpha, A) = r_1$ und damit $r_1 \cap r_2 = \{P^\alpha\}$. Wäre $H \in r_1 \cap r_2$, $H \neq P^\alpha$, dann wäre $r_1 = i(P^\alpha, H) = r_2$ und schließlich $A, B \in r_1$. A und B sind aber verbindbar $(A \nparallel B)$.

Zu c). Sei $r \in R$, $r \neq i_1, i_2$ und $A \in r \setminus \{i_1 \cup i_2\}$. Dann ist $A \nparallel P$, und folglich gibt es eine Erzeugende β durch A und P. Nach (1) gibt es auf β einen dritten Punkt B. Nach (7) existiert eine Erzeugende α durch B mit $\alpha|\beta$ und $A^\alpha = P$. Da $A \neq P$ ist, gilt $A\,\mathrm{I}\,\alpha$, und aus A5 folgt, daß genau zwei Punkte P_1, P_2 auf α existieren, die mit A

[6] Mit $|M|$ bezeichnen wir die Kardinalzahl der Menge M.

unverbindbar sind und folglich auch mit $P=A^\alpha$. O. B. d. A. sei $P_1 \in i_1$ und $P_2 \in i_2$. Wir setzen $i_{P_1}=i(P_1, A)$ und $i_{P_2}=i(P_2, A)$.

Nach (10) ist $r=i_{P_1}$ oder $r=i_{P_2}$.

Es gilt $R_1 \cap R_2=\emptyset$: Für ein $r \in R_1 \cap R_2$ gilt nach (10) $r \neq i_1, i_2$, und somit ist $r \cap i_1 \neq \emptyset$ (wegen $r \in R_2$) ein Widerspruch zu (11) a), da danach $r \cap i_1=\emptyset$ (wegen $r, i_1 \in R_1$). ∎

Nach (11) ist die Definition von R_1 und R_2 unabhängig von der Wahl des Punktes P.

Mit Hilfe von R_1 und R_2 definieren wir zwei Relationen auf $\mathscr{P}$

$$A /\!/_+ B :\Leftrightarrow i(A, B) \in R_1 \quad \text{oder} \quad A = B,$$
$$A /\!/_- B :\Leftrightarrow i(A, B) \in R_2 \quad \text{oder} \quad A = B.$$

Nach (11) sind die beide Relationen $/\!/_+$ und $/\!/_-$ Äquivalenzrelationen.

Es gilt nach A1, A2, A5, (6), (10) und (11) dann sogar

(12) *Die Gruppenebene* $(\mathscr{P}, \mathfrak{Z})$ *mit* I *als Inzidenzrelation und den Äquivalenzrelationen* $/\!/_+$ *und* $/\!/_-$ *auf* $\mathscr{P}$ *ist eine Minkowski-Ebene.*

Nach (12) und 1.3 gilt

(13) *Die von einer Erzeugenden* α *der Gruppe* G *induzierte Abbildung*

$$\bar{\alpha}: \mathscr{P} \to \mathscr{P} \quad \text{mit} \quad \bar{\alpha}(A) = A^\alpha$$

ist die Zykelspiegelung an dem Zykel α *der Gruppenebene.* ∎

(14) *Ist* α *eine Erzeugende und* A *ein Punkt mit* $A \not{\mathrm{I}} \alpha$, *so gilt für jede Erzeugende* β *mit* $A, A^\alpha \mathrm{I} \beta$ *stets* $\beta | \alpha$.

Beweis. Es seien α, A mit $A \not{\mathrm{I}} \alpha$ gegeben. Dann ist $A^\alpha \neq A$. Sei β eine Erzeugende mit $A, A^\alpha \mathrm{I} \beta$. Nach (1) gibt es einen Punkt B, so daß $B \mathrm{I} \beta$ und B ist mit A (und folglich mit A^α) verbindbar. Nach A8 gibt es eine Erzeugende β' mit $A, B \mathrm{I} \beta'$ und $\beta' | \alpha$. Wegen $A \mathrm{I} \beta'$ ist $(A^\alpha)^{\beta'} = A^{\alpha\beta'} = A^{\beta'\alpha} = A^\alpha$, d. h. $A^\alpha \mathrm{I} \beta'$. Nach A2 ist damit aber $\beta' = \beta$ und schließlich $\alpha | \beta$. ∎

Aus (12), (13) und (14) folgt

(15) *Die Gruppenebene* $(\mathscr{P}, \mathfrak{Z})$ *ist eine Minkowski-Ebene, in der an jedem Zykel eine Spiegelung existiert und die den Eigenschaften* (ii), (iii) *und* (iv) *aus* (15) *in* 1.3 *genügt.* ∎

2.3. *Ein Dreispiegelungssatz und die Gruppenebene als miquelsche Minkowski-Ebene*

Wir nutzen jetzt den Satz von Dienst [3], vgl. auch 1.4, um nachzuweisen, daß die auf der Grundlage des gruppentheoretischen Axiomensystems in 2.1 eingeführte Gruppenebene $(\mathscr{P}, \mathfrak{Z})$ eine miquelsche Minkowski-Ebene ist.

Für die nachfolgenden abbildungstheoretischen Betrachtungen identifizieren wir häufig eine Erzeugende α mit der Menge $\{A \in \mathscr{P}: A \mathrm{I} \alpha\}$ der mit α inzidierenden Punkte. Wir verwenden im weiteren auch einige Vereinbarungen und Bezeichnungen aus dem Abschnitt 1, wie z. B. P^+, P^-, $\bar{P}$. Nach (15) ist die Gruppenebene

$(\mathscr{P}, \mathfrak{Z})$ eine S-Minkowski-Ebene. Damit können wir uns nun auf die Sätze (16), (17) und (18) in 1.3 stützen. Satz 16 besagt, daß für alle Erzeugenden α, β mit $A\mathrm{I}\alpha$, β und $\alpha|\beta$ entweder $\alpha\cap\beta=\{A\}$ oder $\alpha\cap\beta=\{A, B\}$ für einen von A verschiedenen Punkt B gilt.

Die Sätze (17) und (18) aus 1.3 präzisieren diese Fallunterscheidung mittels Eigenschaften über den affinen Abschluß $\bar{\mathfrak{A}}_X$ in jedem Punkt der Ebene: $\bar{\mathfrak{A}}_X$ ist für alle X eine Translationsebene, und zwar entweder von einer Charakteristik $\neq 2$ oder von Charakteristik 2.

Wir betrachten nun zwei verbindbare Punkte P, Q, die Menge aller durch P, Q gehenden Erzeugenden und die eindeutig bestimmten Punkte R, S mit $P//_+R//_-Q$ und $P//_-S//_+Q$. Es sei $H:=\{\bar{\alpha}: P, Q\mathrm{I}\alpha\}$ und $\mathscr{G}$ die von H erzeugte Gruppe.

(16) *Es sei A ein beliebiger Punkt und $\mathscr{G}(A)$ die Bahn von A bezüglich $\mathscr{G}$. Dann gilt*
- a) $\mathscr{G}(A)=\{P\}$, *wenn* $A=P$ *ist.*
- b) $\mathscr{G}(A)=\{Q\}$, *wenn* $A=Q$ *ist.*
- c) $\mathscr{G}(A)=\{R, S\}$, *wenn* $A=R$ *oder* $A=S$ *ist.*
- d) $\mathscr{G}(A)=\{X: (X//_+P \vee X//_-P)\wedge X\neq R, S\}=\bar{P}\setminus\{R, S\}$ *wenn* $P//A \not\Vdash Q$ *gilt.*
- e) $\mathscr{G}(A)=\{X: (X//_+Q \vee X//_-Q)\wedge X\neq R, S\}=\bar{Q}\setminus\{R, S\}$ *wenn* $Q//A \not\Vdash P$ *gilt.*
- f) $\mathscr{G}(A)=\alpha\setminus\{R, S\}$ *mit* $\alpha=(A, R, S)$, *falls* $A \not\Vdash R, S$ *gilt.*

Beweis. Die Behauptungen a) und b) sind evident.

Zu c). Es sei o. B. d. A. $A=R$. Dann gilt für alle $\bar{\alpha}\in H$ stets $R^\alpha=S$, da $P, Q\mathrm{I}\alpha$, $R\not\mathrm{I}\alpha$ und R, S nach A4 die einzigen mit P, Q unverbindbare Punkte sind.

Zu d). Es sei $P//A \not\Vdash Q$. O. B. d. A. gelte $A//_-P$. Für jede Erzeugende α mit $\bar{\alpha}\in H$ gilt dann $A^\alpha\in P^+$, da $A\in P^-$, $A^\alpha//P$ und A, A^α nach (5) verbindbar sind. Es sei T ein Punkt mit $T\in P^+$ und $T\neq R$. Sei $\{L\}=A^+\cap T^-$. Es gilt $L \not\Vdash P, Q$, und folglich gibt es eine Erzeugende α_0 mit $L, P, Q\mathrm{I}\alpha_0$. Es ist $\bar{\alpha}_0\in H$ und $A^{\alpha_0}=T$. Ist $V\in P^-$ mit $V\neq S$, so betrachten wir A^α für ein α mit $\bar{\alpha}\in H$. Es sei weiterhin α' die Erzeugende mit $P, Q, K\mathrm{I}\alpha'$, wobei $\{K\}=(A^\alpha)^-\cap V^+$ ist. Es gilt $(A^\alpha)^{\alpha'}=V$.

Zu e). Genauso wie d).

Zu f). Zunächst sei bemerkt, daß nach (8) für jede Erzeugende λ mit $P^\lambda=Q$ gilt $R, S\mathrm{I}\lambda$. Offensichtlich folgt aus $P^\lambda=Q$ wegen $P\neq Q$ stets $P, Q\not\mathrm{I}\lambda$. Nach (14) gilt für jede Erzeugende λ mit $P^\lambda=Q$ und jede Erzeugende α mit $P, Q\mathrm{I}\alpha$ stets $\lambda|\alpha$, d. h., $\bar{\lambda}$ zentralisiert H und folglich die davon erzeugte Gruppe $\mathscr{G}$. Also bleibt jede Erzeugende λ mit $P^\lambda=Q$ unter $\mathscr{G}$ invariant. Nach unseren einleitenden Bemerkung können wir zwei Klassen von Gruppenebenen unterscheiden.

Fall 1. Zwei in der |-Relation stehende Erzeugende mit einem gemeinsamen Punkt berühren sich stets. Daraus folgt, daß für jedes α mit $\bar{\alpha}\in H$ und jedes λ mit $P^\lambda=Q$ $|\alpha\cap\lambda|$ höchstens gleich 1 ist.

Sind $A, B\,\mathrm{I}\,\lambda$ mit $A, B \neq R, S$ und ist α eine Erzeugende mit $Q\,\mathrm{I}\,\alpha$, die die Erzeugende μ mit $A, B, P\,\mathrm{I}\,\mu$ im Punkt P berührt, so vertauscht $\bar{\alpha}\in H$ die Punkte A und B. In der Tat, nach (18) in 1.3 ist $\mu\alpha$ involutorisch, d. h. $\mu\alpha=\alpha\mu$, oder damit äquivalent $\mu^{\alpha}=\mu$. Für λ und α ist schon oben $\lambda|\alpha$ bewiesen worden. Es ist $\{A, B\}^{\alpha}= =(\mu\cap\lambda)^{\alpha}=\mu\cap\lambda=\{A, B\}$, und wegen $A, B\not{\mathrm{I}}\,\alpha$ gilt dann $A^{\alpha}=B$ und $B^{\alpha}=A$. Folglich wirkt die Gruppe $\mathcal{G}$ transitiv auf $\lambda\setminus\{R, S\}$. Mit der Invarianz von $\lambda\setminus\{R, S\}$ unter $\mathcal{G}$ ist somit $\lambda\setminus\{R, S\}$ eine Bahn von $\mathcal{G}$.

Fall 2. Für alle Erzeugenden α, β mit $\alpha|\beta$ und $A_1\,\mathrm{I}\,\alpha, \beta$ gibt es einen Punkt $A_2\neq A_1$ mit $A_2\,\mathrm{I}\,\alpha, \beta$.

Sind A, B verschiedene Punkte auf $\lambda\setminus\{R, S\}$, so sind A, B, P drei paarweise verbindbare Punkte und folglich gibt es genau eine Erzeugende μ mit $A, B, P\,\mathrm{I}\,\mu$. Ist $Q\,\mathrm{I}\,\mu$, so sei ν die nach (19) in 1.3 existierende Erzeugende durch P und Q mit $\nu|\mu$. Dann gilt $\mu^{\nu}=\mu$. Ist $Q\not{\mathrm{I}}\,\mu$, so sind Q, Q^{μ}, P drei paarweise verbindbare Punkte, und es gibt eine Erzeugende ν mit $Q, Q^{\mu}, P\,\mathrm{I}\,\nu$. Nach (14) folgt $\nu|\mu$ und folglich in diesem Fall auch $\mu^{\nu}=\mu$. Wegen $\bar{\nu}\in H$ ist $\lambda^{\nu}=\lambda$ und $\{A, B\}^{\lambda}=(\lambda\cap\mu)^{\nu}=\lambda\cap\mu= =\{A, B\}$. A kann kein Fixpunkt von $\bar{\nu}$ sein, denn sonst wären A, B und P Fixpunkte von $\bar{\nu}$, was $\nu=\mu$ wegen (13) und A2 zur Folge hätte. Dies kann aber nicht sein, weil aus $\nu|\mu$ folgt $\nu\neq\mu$. Also ist $A^{\nu}=B$. Daher ist $\lambda\setminus\{R, S\}$ eine Bahn von $\mathcal{G}$.

Nach (7) bilden die Erzeugenden λ mit $P^{\lambda}=Q$ eine Überdeckung von $\mathcal{P}\setminus(\bar{P}\cup\bar{Q})$. Damit ist die Menge der Bahnen die Menge aller λ' mit $\lambda'=\lambda\setminus\{R, S\}$ und $P^{\lambda}=Q$. ∎

(17) *Berühren sich zwei in der |-Relation stehende Erzeugende, so ist* $H^3\subset H$.

Beweis. H^2 enthält kein Element der Ordnung 2, welches verschieden von der Identität ist. In der Tat, wäre $\bar{\alpha}\bar{\alpha}_0\in H^2$ ein Element der Ordnung 2, dann gelte $\bar{\alpha}\bar{\alpha}_0\bar{\alpha}\bar{\alpha}_0=\mathrm{id}$ und $\alpha\neq\alpha_0$, also $\bar{\alpha}\bar{\alpha}_0=\bar{\alpha}_0\bar{\alpha}$ und folglich $\alpha^{\alpha_0}=\alpha$, d. h. $|\alpha\cap\alpha_0|=1$, was $P, Q\,\mathrm{I}\,\alpha, \alpha_0$ widerspricht.

Sind φ, σ, τ mit $\bar{\varphi}, \bar{\sigma}, \bar{\tau}\in H$ und ist λ die Bahn von $\mathcal{G}$ durch den Punkt $T\in\tau\setminus\{P, Q\}$, so existiert ein $\bar{\alpha}\in H$ mit $T^{\sigma\varphi\alpha}=T$. Die Fixerzeugende τ von $\bar{\sigma}\bar{\varphi}\bar{\alpha}$ trifft jede Bahn in höchstens einem Punkt und besteht nur aus Fixpunkten, d. h., es ist $\bar{\sigma}\bar{\varphi}\bar{\alpha}=\bar{\tau}$ oder $\bar{\sigma}\bar{\varphi}\bar{\alpha}=\mathrm{id}_H$. Die 2. Möglichkeit kann nicht eintreten, da H^2 keine Involution enthält. Daher ist $\bar{\tau}\bar{\sigma}\bar{\varphi}=\bar{\alpha}$, d. h., $H^3\subset H$. ∎

(18) *Es sei der affine Abschluß* $\bar{\mathfrak{A}}_X$ *für einen Punkt* X *der Gruppenebene eine Translationsebene einer Charakteristik* $\neq 2$. *Ist dann* $A\in\mathcal{P}\setminus(\bar{P}\cup\bar{Q})$ *und* β *die Erzeugende mit* $P, Q, A\,\mathrm{I}\,\beta$, *so besteht die Standgruppe* $\mathcal{G}_A:=\{a\in\mathcal{G}: A^a=A\}$ *aus* $\mathrm{id}_{\mathcal{P}}$ *und* $\bar{\beta}$.

Beweis. Es ist klar, daß $\mathrm{id}_{\mathcal{P}}$ und $\bar{\beta}$ Elemente der Standgruppe sind. Es sei $a\in\mathcal{G}_A$. $a|_{\mathcal{P}\setminus P}$ ist eine Kollineation von $\bar{\mathfrak{A}}_P$ mit den Fixpunkten Q, A und der Fixgeraden $\beta\setminus\{P\}$. Für jeden Punkt $X\in\beta\setminus\{P\}$ bezeichnet t_X jene Translation der Translationsebene $\bar{\mathfrak{A}}_P$, die Q in X abbildet. Es ist

$$U:=\{t_X: X\in\beta\setminus\{P\}\}$$

eine Untergruppe der Translationsgruppe von $\bar{\mathfrak{A}}_P$.

Ist λ' die Bahn von $X \in \beta \setminus \{P, Q\}$ unter $\mathscr{G}$, so ist λ' wegen (16) eine um die Punkte R, S verminderte Erzeugende λ mit $P^\lambda = Q$, und es gibt wegen (17) in 1.3 einen Punkt X', so daß $\lambda \cap \beta = \{X, X'\}$. Da $\beta^a = \beta$ ist, gilt $X^a \in \beta \cap \lambda = \{X, X'\}$ und

(I) $$t_X^a := a^{-1} t_X a = t_X a \in \{t_X, t_{X'}\} \quad \text{(vgl. Bild)}$$

$$\begin{tikzcd} Q \arrow[r, "t_X"] \arrow[d, "a"'] & X \arrow[d, "a"] \\ Q \arrow[r, "t_X^a = t_X a"] & X^a \end{tikzcd}$$

Wir beweisen jetzt, daß $t_{X'} = t_X^{-1}$ ist. Es sei π die Spiegelung an Q in $\mathfrak{A}_P$. Dann ist $X' = X^\pi$. Sei π' die Spiegelung an M, dem Mittelpunkt von Q und X. Dann ist $t_X = \pi\pi'$ (vgl. Pickert [8], S. 211—213) und folglich

$$t_X \pi = t_X^\pi = \pi t_X \pi = \pi(\pi\pi')\pi = \pi'\pi = t_X^{-1}.$$

Dann ist

$$t_X^a = t_X a \in \{t_X, t_X^{-1}\} \quad \text{(vgl. (I))}.$$

Wir betrachten die Konjugation

$$a^*: U \to U \quad \text{mit} \quad t_X^{a^*} = t_X^a = a^{-1} t_X a = t_X a \in \{t_X, t_X^{-1}\}.$$

Da $\beta^a = \beta$ ist, folgt aus $X \in \beta$ stets $X^a \in \beta$ und somit $t_X a \in U$, d. h. a^* ist ein Automorphismus von U. Wir haben $t_A \in U$, $t_A \neq \mathrm{id}$ und $t_A^2 \neq \mathrm{id}$, da die Charakteristik der Translationsebene $\neq 2$ ist.

Es gilt $t_A^{a^*} = a^{-1} t_A a = t_A a = t_A$ wegen $A^a = A$.

Wir beweisen jetzt, daß $a^* = \mathrm{id}_U$ ist. Angenommen, es gibt ein $X \in \beta \setminus \{P\}$ mit $X \neq Q$, so daß $t_X^{a^*} = t_X^{-1}$. Da U eine Gruppe ist, existiert dann eine Translation $t_Y \in U$ mit $t_A = t_Y t_X$. Dann haben wir

$$t_Y t_X = t_A = t_A^{a^*} = (t_Y t_X)^{a^*} = t_Y^{a^*} t_X^{a^*} = t_Y^{a^*} t_X^{-1}.$$

Es gilt $t_Y^{a^*} \in \{t_Y, t_Y^{-1}\}$.

Fall 1. $t_Y^{a^*} = t_Y$. Aus $t_A = t_Y t_X = t_Y t_X^{-1}$ folgt $t_X = t_X^{-1}$. Aber das ist ein Widerspruch, da t_X nicht involutorisch ist.

Fall 2. $t_Y^{a^*} = t_Y^{-1}$. Es gilt hier $t_A = t_Y t_X = t_Y^{-1} t_X^{-1} = t_A^{-1}$. Auch das ist ein Widerspruch zu der Tatsache, daß t_A nicht involutorisch ist.

Folglich ist $a^* = \mathrm{id}_U$.

Die Abbildung a läßt, wegen $t_X^{a^*} = t_X^a = t_X a$ jeden Punkt $X \in \beta \setminus \{P\}$ fest. Da auch P ein Fixpunkt von a ist, ist $a \in \{\mathrm{id}_{\mathscr{P}}, \beta\}$. ■

Wir können jetzt den Dreispiegelungssatz für Erzeugende durch zwei Punkte P, Q allgemein beweisen.

(19) *In der Gruppenebene* $(\mathscr{P}, \mathfrak{Z})$ *gilt stets* $H^3 \subset H$.

Beweis. Falls sich zwei in der |-Relation stehende Erzeugende berühren, gilt $H^3 \subset H$ nach (17). Anderfalls ist der affine Abschluß $\mathfrak{A}_X$ für einen Punkt X der Gruppenebene eine Translationsebene einer Charakteristik $\neq 2$. Es gilt auch hier zunächst $H \cap H^2 = \emptyset$: Wäre $a \in H \cap H^2$, so gäbe es zwei Erzeugende α, α' mit $P, Q \mathrm{I} \alpha, \alpha'$ und $a = \bar{\alpha}\bar{\alpha}' = \bar{\alpha}'\bar{\alpha}$. Dann wäre $a|_{\mathscr{P} \setminus P}$ die Punktspiegelung in $\mathfrak{A}_P$ am Punkt Q, d. h., a hätte nur die Fixpunkte P, Q. Als Abbildung von H müßte a jedoch eine ganze Erzeugende punktweise festlassen, also mindestens drei Fixpunkte besitzen.

Es seien $\bar{\alpha}, \bar{\beta}, \bar{\gamma} \in H$ und $A \in \alpha \setminus \{P, Q\}$. Nach (16) existiert eine Abbildung $\bar{\delta} \in H$ mit $A^{\bar{\beta}\bar{\gamma}\bar{\delta}} = A$. Mit (18) folgt $\bar{\beta}\bar{\gamma}\bar{\delta} \in \mathscr{G}_A = \{\mathrm{id}_{\mathscr{P}}, \bar{\alpha}\}$. Wegen $H^2 \cap H = \emptyset$ kann nicht $\bar{\beta}\bar{\gamma}\bar{\delta} = \mathrm{id}_{\mathscr{P}}$ sein, also muß $\bar{\beta}\bar{\gamma}\bar{\delta} = \bar{\alpha}$ bzw. $\bar{\delta} = \bar{\alpha}\bar{\beta}\bar{\gamma}$ gelten. ∎

Aus (15), (17) und (19) folgt

SATZ 1. *Ist* G *eine Gruppe, für die die Grundannahme zutrifft und die Axiome* A1 *bis* A8 *gelten, so ist die Gruppenebene* $(\mathscr{P}, \mathfrak{Z})$ *eine miquelsche Minkowski-Ebene.*

2.4 *Die von Zykelspiegelungen erzeugte Gruppe einer miquelschen Minkowski-Ebene*

Nachdem in 2.1 bis 2.3 aus Annahmen über eine Gruppe in der Gruppe ein geometrische Struktur eingeführt werden konnte, die sich als miquelsche Minkowski-Ebene herausstellte, wollen wir nun zeigen, daß diese Annahmen für jede von Zykelspiegelungen erzeugte Gruppe einer beliebigen miquelschen Minkowski-Ebene zutreffen.

SATZ 2. *Ist* $M = (\mathfrak{P}, \mathfrak{Z})$ *eine miquelsche Minkowski-Ebene, so genügt die von den Spiegelungen an Zykeln erzeugte Gruppe* $\mathscr{G}(M)$ *der Grundannahme und den Axiomen* A1 *bis* A8.

BEWEIS. Es sei $A \in \mathfrak{P}$. Wir betrachten die Mengen

$$\hat{A} := \{\sigma_z : \sigma_z(A) = A\}.$$

Sei

$$\mathscr{P} := \{\hat{A} : A \in \mathfrak{P}\}.$$

Zu (I). a) Aus $\sigma_z \in \hat{A}$ folgt $\sigma_z \hat{A} \sigma_z = \hat{A}$, da nach (13) in 1.3 für

$$\sigma_{\mathfrak{z}} \in \hat{A} \quad \text{gilt} \quad \sigma_z \sigma_{\mathfrak{z}} \sigma_z = \sigma_{\sigma_z(\mathfrak{z})} \in \hat{A}.$$

b) Es sei $\sigma_{\mathfrak{z}} \in \hat{A}$ mit $\sigma_z \sigma_{\mathfrak{z}} \sigma_z \in \hat{A}$. Dann folgt aus $\sigma_z \sigma_{\mathfrak{z}} \sigma_z = \sigma_{\sigma_z(\mathfrak{z})} \in \hat{A}$ mit (8) aus 1.3 stets $A \in \sigma_z(\mathfrak{z})$ bzw. $\sigma_z(A) \in \mathfrak{z}$.

Wäre $\sigma_z(A) \neq A$, so müßte für ein existierendes z' durch A $(\sigma_{z'}(A) = A)$ mit $z' \neq z$ und $\sigma_z(A) \notin z'$ (vgl. dazu das Minimalmodell, das aus 9 Punkten und 6 Zykeln besteht) aus $\sigma_z \sigma_{z'} \sigma_z \in \hat{A}$ wie oben der Widerspruch $\sigma_z(A) \in z'$ folgen.

Somit ist $\sigma_z(A) = A$, d. h. $\sigma_z \in \hat{A}$.

Zu (II). Es sei σ_z eine Spiegelung an z und $\hat{A} \in \mathscr{P}$. Für alle $\sigma_{\mathfrak{z}} \in \hat{A}$ gilt (nach (8) in 1.3) $A \in \mathfrak{z}$, und es sei $A' = \sigma_z(A)$. Nach (13) in 1.3 besteht $\sigma_z \hat{A} \sigma_z$ nur aus Spiegelungen, und jede solche Spiegelung ist eine an einem Zykel durch A'. Da σ_z involutorisch ist, werden alle Zykel durch A' erfaßt, d. h., es gilt insgesamt $\sigma_z \hat{A} \sigma_z = \hat{A}'$.

Zu (III). Es sei σ_z eine beliebige Zykelspiegelung. Da jeder Zykel mindestens drei Punkte enthält, gibt es einen Punkt A mit $A \in z$. Daraus folgt $\sigma_z(A) = A$, d. h. $\sigma_z \in \hat{A}$.

Zu A1. Nach M4 gibt es drei paarweise nicht parallele Punkte A, B, C. Nach (5) in 1.1 geht durch A und B mindestens ein Zykel z. Folglich ist $\sigma_z \in \hat{A}, \hat{B}$ d. h. $\hat{A}$ und $\hat{B}$ sind verbindbar. Genauso folgt dies für $\hat{A}$, $\hat{C}$ und $\hat{B}$, $\hat{C}$.

Zu A2. Nach M1 geht durch drei paarweise nicht parallele Punkte A, B, C genau ein Zykel z. Dann ist σ_z die einzige Zykelspiegelung mit $\sigma_z \in \hat{A}, \hat{B}, \hat{C}$, d. h., es gilt A2.

Zu A3. Es seien σ_z und $\hat{A}$ gegeben mit $\sigma_z \in \hat{A}$. Da jeder Zykel mindestens drei Punkte enthält, seien $C, D \in z$, $C, D \neq A, C \neq D$; und sei $\{B\} = C^+ \cap D^-$. Es gilt $B \notin z$, und B ist nicht zu A parallel. Dann gibt es nach (5) in 1.1 einen Zykel z' mit $A, B \in z'$, d. h., es existiert $\hat{B}$ mit $\sigma_z \notin \hat{B}$ und $\sigma_{z'} \in \hat{A}, \hat{B}$. Folglich gilt A3.

Zu A4. Es seien $\hat{A}$, $\hat{B}$ zwei verbindbare Punkte, d. h., es existiert $z \in \mathfrak{Z}$, so daß $\sigma_z \in \hat{A}, \hat{B}$. Es seien $\{C\} := A^+ \cap B^-$ und $\{D\} := A^- \cap B^+$. Die Punkte, die zu B parallel sind, gehören zu B^+ oder zu B^- und die Punkte, die zu A parallel sind, gehören zu A^+ oder zu A^-. Nach M3 trifft A^+ (bzw. A^-) die Erzeugende B^- (bzw. B^+) in genau einem Punkt. Dann sind C und D die einzigen Punkte, die zu A und B parallel sind. Offensichtlich sind dann $\hat{C}$ und $\hat{D}$ genau die Elemente von $\mathscr{P}$, die weder mit $\hat{A}$ noch mit $\hat{B}$ verbindbar sind.

Zu A5. Folgt unmittelbar aus M2 und (1) in 1.1.

Zu A6. Es seien σ_z, $\sigma_{z'}$, $\sigma_{z''}$ drei Spiegelungen an Zykeln und $\hat{A}, \hat{B} \in \mathscr{P}$ mit $\hat{A} \neq \hat{B}$, $\sigma_z, \sigma_{z'} \in \hat{A}, \hat{B}$ und $\sigma_{z''} \in \hat{A}$. Dann sind $A, B \in \mathfrak{P}$, mit $A, B \in z, z'$ und $A \in z''$. Betrachten wir die affine Ebene $\mathfrak{A}_A$. In $\mathfrak{A}_A$ sind $z \setminus \{A\}$, $z' \setminus \{A\}$, $z'' \setminus \{A\}$ drei Geraden, und $z \setminus \{A\}$ und $z' \setminus \{A\}$ schneiden sich in B. Folglich ist $(z'' \setminus \{A\}) \cap \cap (z \setminus \{A\}) \neq \emptyset$ oder $(z'' \setminus \{A\}) \cap (z' \setminus \{A\}) \neq \emptyset$, und damit gibt es einen von A verschiedenen Punkt $C \in z''$ mit $C \in z$ oder $C \in z'$, d. h. $\sigma_z(C) = C$ oder $\sigma_{z'}(C) = C$ und damit $\hat{C} \mathrm{I} \sigma_z$ oder $\hat{C} \mathrm{I} \sigma_{z'}$.

Zu A7. Es seien $\hat{A}$, $\hat{B}$ zwei Punkte mit $\sigma_z \sigma_{z'} \sigma_{z''} \in \hat{A}, \hat{B}$, und $\sigma_z | \sigma_{z'}, \sigma_{z''}$. Dann sind $A, B \in \mathfrak{P}$ mit $A, B \in z, z', z''$ und $z \perp z', z''$. Wir betrachten die affine Ebene $\mathfrak{A}_A$. Dann sind $z \setminus \{A\}$, $z' \setminus \{A\}$, $z'' \setminus \{A\}$ drei Geraden durch B in $\mathfrak{A}_A$ mit $(z \setminus \{A\}) \perp (z' \setminus \{A\}), (z'' \setminus \{A\})$. Daraus folgt $(z' \setminus \{A\}) = (z'' \setminus \{A\})$ und folglich $z' = z''$ und damit $\sigma_{z'} = \sigma_{z''}$.

Zu A8. Es seien σ_z eine Spiegelung an $z \in \mathfrak{Z}$ und $\hat{A}, \hat{B} \in \mathscr{P}$ mit $\hat{A} \neq \hat{B}$; $\sigma_z \notin \hat{A}$ und $\hat{B}$ mit $\hat{A}$ und $\widehat{\sigma_z(A)}$ verbindbar. Dann sind A, $\sigma_z(A)$, B drei paarweise nicht parallele Punkte und $A \notin z$ gilt. Da M eine miquelsche Minkowski-Ebene ist, gilt die Aussage (iv) in (15) aus 1.3 (vgl. Dienst [3] S. 205), d. h., es gibt einen Zykel z' mit $A, B \in z'$ und $z \perp z'$. Daraus folgt $\sigma_z \sigma_{z'} = \sigma_{z'} \sigma_z \neq \mathrm{id}$, d. h. $\sigma_z \sigma_{z'}$ ist involutorisch und $\sigma_{z'} \in \hat{A}, \hat{B}$. ∎

2.5. *Isomorphien*

Satz 3. *Es sei M eine miquelsche Minkowski-Ebene und $\mathscr{G}(M)$ die von den Spiegelungen an Zykeln von M erzeugte Gruppe. (So genügt $\mathscr{G}(M)$ der Grundannahme und den Axiomen* A1 *bis* A8.) *Es sei $\mathscr{M}(\mathscr{G}(M))$ die Gruppenebene von $\mathscr{G}(M)$. Dann ist M isomorph zu $\mathscr{M}(\mathscr{G}(M))$.*

Beweis. Wir betrachten die Abbildung

$$\mathfrak{i}: M \to \mathcal{M}(\mathcal{G}(M))$$

mit $\mathfrak{i}(P)=\{\sigma_z: \sigma_z(P)=P$ und $\mathfrak{i}(z)=\sigma_z$. Offenbar ist $\mathfrak{i}$ bijektiv und aus $P\in z$ folgt $\sigma_z(P)=P$, d. h. $\mathfrak{i}(z)\in\mathfrak{i}(P)$ und damit ist $\mathfrak{i}$ auch inzidenztreu.

Wir beweisen jetzt, daß

$$P/\!/_+ Q \Rightarrow \mathfrak{i}(P)/\!/_+\mathfrak{i}(Q), \quad (P \neq Q),$$

$$R/\!/_- P \Rightarrow \mathfrak{i}(R)/\!/_-\mathfrak{i}(P), \quad (R \neq P).$$

Da $P/\!/_+ Q$, existiert kein Zykel durch P und Q, d. h.

$$\{\sigma_z:\sigma_z(P) = P\} \cap \{\sigma_z:\sigma_z(Q) = Q\} = \emptyset$$

und folglich ist $\mathfrak{i}(P)/\!/\mathfrak{i}(Q)$. Genauso beweist man $\mathfrak{i}(R)/\!/\mathfrak{i}(P)$. Aus $R/\!/_- P/\!/_+ Q$, $R\neq P\neq Q$ folgt, daß R und Q nicht parallel sind, und folglich ist $\mathfrak{i}(Q)/\!/\mathfrak{i}(R)$. Dann sind $i(\mathfrak{i}(P), \mathfrak{i}(Q))$ und $i(\mathfrak{i}(P), \mathfrak{i}(R))$ die isotropen Geraden durch $\mathfrak{i}(P)$, d. h. $\mathfrak{i}(P)/\!/_+\mathfrak{i}(Q)$ und $\mathfrak{i}(P)/\!/_-\mathfrak{i}(R)$. (Die Benennung + oder − spielt hier keine Rolle.) ∎

Satz 4. *Es sei* **G** *eine Gruppe, die der Grundannahme und den Axiomen* A1 *bis* A8 *genügt. (Dann ist die Gruppenebene* $\mathcal{M}$(**G**) *eine miquelsche Minkowski-Ebene.) Es sei* $\mathcal{G}(\mathcal{M}(\mathbf{G}))$ *die von den Spiegelungen an Zykeln von* $\mathcal{M}$(**G**) *erzeugte Gruppe. Dann ist* **G** *isomorph zu* $\mathcal{G}(\mathcal{M}(\mathbf{G}))$.

Beweis. Wir betrachten die Abbildung

$$\mathfrak{i}:\mathbf{G} \to \mathcal{G}(\mathcal{M}(\mathbf{G})) \quad \text{mit} \quad \mathfrak{i}(a) = \bar{a},$$

wo

$$\bar{a}:\mathcal{P} \to \mathcal{P}, \quad \text{so daß} \quad \bar{a}(A) = A^a.$$

Nach (13) in 2.2 ist $\bar{a}$ eine bijektive Abbildung. Wir beweisen, daß $\mathfrak{i}$ ein Homomorphismus ist.

$\mathfrak{i}(ab)=\overline{ab}$. Für alle $A\in\mathcal{P}$ gilt $ab(A)=A^{ab}=(A^a)^b=\bar{b}(A^a)=\bar{b}(\bar{a}(A))$ d. h. $\mathfrak{i}(ab)=\mathfrak{i}(a)\mathfrak{i}(b)$. Da aus $A^a=A$ für alle Punkte A aus $\mathcal{M}$(**G**) folgt, daß a das Einselement der Gruppe **G** ist (vgl. (12) in 1.1 und (13) in 2.2), und damit ist $\mathfrak{i}$ sogar ein Isomorphismus. ∎

Auf Grund der Sätze 1 und 2 heben wir zusammenfassend hervor:

Satz 5. *Eine Gruppe* **G** *ist genau dann die von den Zykelspiegelungen einer miquelschen Minkowski-Ebene erzeugten Automorphismengruppe, wenn sie der Grundannahme und den Axiome* A1 *bis* A8 *genügt.*

LITERATURVERZEICHNIS

[1] Bachmann, F., *Aufbau der Geometrie aus dem Spiegelungsbegriff*, 2. Aufl., Die Grundlehren der mathematischen Wissenschaften, Band 96, Springer-Verlag, Berlin—Heidelberg—New York, 1973. *MR* 49 # 11368.

[2] Benz, W., *Vorlesungen über Geometrie der Algebren.* Geometrien von Möbius, Laguerre—Lie, Minkowski in einheitlicher und grundlagengeometrischer Behandlung, Die Grund-

lehren der mathematischen Wissenschaften, Band 197, Springer-Verlag, Berlin—Heidelberg—New York, 1973. *MR* **50** # 5623.

[3] Dienst, K. J., Minkowski-Ebenen mit Spiegelungen, *Monatsh. Math.* **84** (1977), 197—208. *MR* **58** # 23999.

[4] Karzel, H. und Mäurer, H., Eine Kennzeichnung miquelscher Möbius-Ebenen durch eine Eigenschaft der Kreisspiegelungen, *Resultate Math.* **5** (1982), 52—56. *MR* **83h:** 51016.

[5] Lang, K., Kennzeichnung von Möbius-Ebenen durch Spiegelungsrelationen, *J. Geom.* **16** (1981), 5—18. *MR* **82j:** 51009.

[6] Lang, K., Zur Kennzeichnung miquelscher Minkowski-Ebenen, *J. Geom.* **18** (1982), 99—108. *MR* **83k:** 51012.

[7] Mäurer, H., Kreisspiegelungen in Möbius-Ebenen, *Geometriae Dedicata* **2** (1973), 261—268. *MR* **52** # 6418.

[8] Molnár, E., Begründung der Möbiusschen Kreisebene aus dem Spiegelungsbegriff, *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.* **18** (1975), 143—170. *MR* **55** # 6303.

[9] Pickert, G., *Projektive Ebenen,* 2. Aufl., Die Grundlehren der mathematischen Wissenschaften, Band 80, Springer-Verlag, Berlin—Heidelberg—New York, 1975. *MR* **51** # 6577.

[10] Wernicke, B., Möbius-Ebenen in spiegelungsgeometrischer Darstellung, *Studia Sci. Math. Hungar.* **21** (1986), 363—372.

[11] Wernicke, B., On reflections in Möbius planes, *Intuitive Geometry,* Colloquia Math. Soc. J. Bolyai, Vol. 48, North-Holland Publ. Co., Amsterdam—Oxford—New York, 1987, 617—629.

[12] Zeitler, H., On reflections in Minkowski planes, *Geometry and differential geometry* (Proc. conf., Univ. Haifa, Haifa, 1979), Lecture Notes in Mathematics, vol. 792, Springer-Verlag, Berlin, 1980, 183—202. *MR* **81m:** 51011.

(Eingegangen am 16. Dezember 1985)

PÄDAGOGISCHE HOCHSCHULE ERFURT/MÜHLHAUSEN
SEKTION MATHEMATIK/PHYSIK
NORDHÄUSER STRASSE 63
DDR—5064 ERFURT
GERMAN DEMOCRATIC REPUBLIC

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 33—36

COMMUTATIVITY OF CERTAIN SEMIPRIME RINGS

HAZAR ABU-KHUZAM and ADIL YAQUB

A theorem of Herstein [2] states that a ring R which satisfies the identity $(xy)^n = x^n y^n$, where n is a fixed positive integer greater than 1, must have nil commutator ideal. In [1], it was proved that a semiprime ring in which for each x in R there exists a positive integer $n=n(x)>1$ such that $(xy)^n = x^n y^n$ for all y in R must be commutative. In this direction we prove the following theorems which generalize the above mentioned results. Throughout Z denotes the center of R.

THEOREM 1. *Let R be a semiprime ring such that for each x in R there exists a positive integer $n=n(x)>1$ such that $[x, (xy)^n - x^n y^n]=0$ and $[x, (yx)^n - y^n x^n]=0$ for all y in R. Then R is commutative.*

THEOREM 2. *Let R be a semiprime ring such that for each x in R there exists a positive integer $n=n(x)>1$ such that $(xy)^n - x^n y^n \in Z$, and $(x^2 y)^n - x^{2n} y^n \in Z$ for all y in R. Then R is commutative.*

In preparation for the proof of Theorem 1, we first prove the following lemmas. Lemma 1 is known and we omit its proof.

LEMMA 1. *If $[x, y]$ commutes with x, then $[x^k, y] = kx^{k-1}[x, y]$ for all positive integers $k>1$.*

LEMMA 2. *If R is a semiprime ring in which, for each x in R, there exists a positive integer $n=n(x)>1$ such that $[x, (xy)^n - x^n y^n]=0$ for all y in R, then R has no nonzero nilpotent elements.*

PROOF. Let $a \in R$ such that $a^2=0$. Using the hypothesis, there exists an integer $n=n(a)>1$ such that $[a, (ay)^n - a^n y^n]=0$ for all y in R. This implies that $(ay)^n a = a(ay)^n = 0$, and hence $(ay)^{n+1}=0$ for all y in R. If $aR \neq 0$, then the above shows that aR is a nonzero nil right ideal satisfying the identity $x^{n+1}=0$ for all x in aR. So by Lemma 2.1.1 of [4], R has a nonzero nilpotent ideal. This is a contradiction since R is semiprime. Thus $aR=0$, and hence $aRa=0$. This implies that $a=0$ since R is semiprime.

LEMMA 3. *If R is a prime ring in which, for each x in R, there exists an integer $n=n(x)>1$ such that $[x, (xy)^n - x^n y^n]=0$ for all y in R, then R has no zero divisors.*

1980 *Mathematics Subject Classification* (1985 Revision). Primary 16A70.
Key words and phrases. Nil ring, commutator ideal, semiprime ring.

PROOF. By Lemma 2 above, R has no nonzero nilpotent elements. So by Lemma 1.1.1 of [4], R has no zero divisors since it is prime with no nonzero nilpotent elements.

PROOF of Theorem 1. Since R is a semiprime ring, then it is isomorphic to a subdirect sum of prime rings R_α each of which, as a homomorphic image of R, satisfies the hypothesis of the Theorem. So we may assume that R is prime. Let x and y be any two nonzero elements of R. Then by the hypothesis, there exists a positive integer $n=n(x)>1$ such that

$$z_1 = (xy)^n - x^n y^n \text{ commutes with } x, \text{ and} \tag{1}$$

$$z_2 = (yx)^n - y^n x^n \text{ commutes with } x. \tag{2}$$

Now $(xy)^n x = x(yx)^n$ and using (1) and (2), this implies that

$$(x^n y^n + z_1)x = x(y^n x^n + z_2).$$

So

$$x^n y^n x - xy^n x^n = x(z_2 - z_1).$$

Thus,

$$x^n y^n x - xy^n x^n = x[x^{n-1}, y^n]x \text{ commutes with } x. \tag{3}$$

Using Lemma 3, R has no zero divisors, and hence (3) implies

$$x[x^{n-1}, y^n] = [x^{n-1}, y^n]x.$$

So

$$[[y^n, x^{n-1}], x^{n-1}] = 0. \tag{4}$$

We now distinguish two cases.

Case 1. Char $R=p\neq 0$. Then using (4) and Lemma 1 we have

$$[y^n, (x^{n-1})^p] = p(x^{n-1})^{p-1}[y^n, x^{n-1}] = 0.$$

Hence R is commutative by a theorem of Herstein [3].

Case 2. Char $R=0$. Then since R is prime, R is torsion-free. So R is a torsion-free domain and hence R is commutative by using (4) and the Lemma of [5]. This completes the proof of Theorem 1.

To prove Theorem 2, we need to prove the following lemma.

LEMMA 4. *Let R be a domain in which for each x in R, there exists a positive integer $n=n(x)>1$ such that either $x^n \in Z$ or $(xy)^n = x^n y^n$ for all y in R. Then R is commutative.*

PROOF. Let x and y be any two elements of R. There exist positive integers $n=n(x)>1$, $m=m(y)>1$ such that

$$\text{either } x^n \in Z \text{ or } (xz)^n = x^n z^n \text{ for all } z \text{ in } R \tag{5}$$

and

$$\text{either } y^m \in Z \text{ or } (yz)^m = y^m z^m \text{ for all } z \text{ in } R. \tag{6}$$

We distinguish two cases.

Case 1. $x^n \notin Z$ and $y^m \notin Z$. Then from (5) and (6) we have

(7) $$(xz)^n = x^n z^n \quad \text{for all } z \text{ in } R$$

and

(8) $$(yz)^m = y^m z^m \quad \text{for all } z \text{ in } R.$$

A simple induction shows that

(9) $$(x^k z)^n = x^{kn} z^n \quad \text{for all } z \text{ in } R \text{ and all positive integers } k,$$

and

(10) $$(y^k z)^m = y^{km} z^m \quad \text{for all } z \text{ in } R \text{ and all positive integers } k.$$

Now, by (9) we have

$$x^k(yx^k)^n = (x^k y)^n x^k = x^{kn} y^n x^k$$

and so

$$x^k\big((yx^k)^n - x^{k(n-1)} y^n x^k\big) = 0.$$

Since R is a domain, this implies that

(11) $$(yx^k)^n = x^{k(n-1)} y^n x^k.$$

Proceeding as done in [1] and using (10) and (11) we have

(12) $$(yx)^{mn} = \big((yx)^m\big)^n = (y^m x^m)^n = x^{m(n-1)} y^{mn} x^m.$$

From (11), $(yx)^n = x^{n-1} y^n x$ and hence

(13) $$\big((yx)^n\big)^m = (x^{n-1} y^n x)^m = x^{n-1} y^n x\, x^{n-1} y^n x \dots x^{n-1} y^n x$$
$$= x^{n-1}(y^n x^n)^{m-1} y^n x.$$

Thus, using (13) and (10),

(14) $$(yx)^{nm} x^{n-1} = [x^{n-1}(y^n x^n)^{m-1} y^n x] x^{n-1} = x^{n-1}(y^n x^n)^m$$
$$= x^{n-1} y^{nm} x^{nm}.$$

Since R has no zero divisors, (14) implies that

(15) $$(yx)^{nm} = x^{n-1} y^{nm} x^{nm-(n-1)}.$$

Now (12) and (15) imply that

(16) $$x^{n-1} y^{nm} x^{nm-(n-1)} = x^{m(n-1)} y^{mn} x^m.$$

Clearly $m(n-1) > n-1$ and $nm-(n-1) = n(m-1)+1 > m$ since $m \geqq 2$, and $n \geqq 2$. So (16) implies that

(17) $$x^{n-1}(y^{nm} x^{nm-m-n+1} - x^{mn-m-n+1} y^{mn}) x^m = 0.$$

Since R has no zero divisors, (17) implies that

(18) $$y^{nm} x^{nm-m-n+1} = x^{mn-m-n+1} y^{mn}.$$

Case 2. Either $x^n \in Z$ or $y^m \in Z$, Then clearly we have

$$y^m x^n = x^n y^m. \tag{19}$$

So in any case, we have from (18) and (19) that $x^p y^q = y^q x^p$ for some positive integers $p = p(x, y)$ and $q = q(x, y)$ and therefore R is commutative by a theorem of Herstein [3].

PROOF of Theorem 2. As done in Theorem 1, we may assume that R is prime. By Lemma 3, R is a domain. Let x be an element of R. Then there exists a positive integer $n = n(x) > 1$ such that $(xy)^n - x^n y^n \in Z$ and $(x^2 y)^n - x^{2n} y^n \in Z$ for all y in R. So there exist elements $z = z(x, y)$, $z' = z'(x, y)$ and $z'' = z''(x, xy)$ in the center Z of R such that

$$(xy)^n = x^n y^n + z, \quad z = z(x, y) \in Z \tag{20}$$

$$(x^2 y)^n = x^{2n} y^n + z', \quad z' = z'(x, y) \in Z \tag{21}$$

$$(x^2 y)^n = [x(xy)]^n = x^n (xy)^n + z'', \quad z'' = z''(x, xy) \in Z. \tag{22}$$

Using (22) and (20)

$$\begin{aligned}(x^2 y)^n = [x(xy)]^n = x^n (xy)^n + z'' &= x^n (x^n y^n + z) + z'' \\ &= x^{2n} y^n + z x^n + z''.\end{aligned} \tag{23}$$

Comparing (21) and (23) we get $zx^n + z'' = z'$ and hence $zx^n \in Z$. Thus

$$zx^n r = zrx^n, \quad z = z(x, y) = (xy)^n - x^n y^n \in Z, \tag{24}$$

for all r in R.

Since we have fixed x, then if for some y in R, $z \neq 0$, (24) implies that $x^n r = rx^n$ for all r in R since R has no zero divisors. So either $z = 0$ for all y in R or $x^n \in Z$. This shows that

$$\text{Either} \quad (xy)^n = x^n y^n \quad \text{for all } y \text{ in } R \text{ or } x^n \in Z, \quad n = n(x). \tag{25}$$

So by Lemma 4 above, R is commutative. This completes the proof of Theorem 2.

REFERENCES

[1] ABU-KHUZAM, H., A commutativity theorem for semiprime rings, *Bull. Austral. Math. Soc.* **27** (1983), 221—224. *MR* **84i:** 16037.

[2] HERSTEIN, I. N., Power maps in rings, *Michigan Math. J.* **8** (1961), 29—32. *MR* **22** #9512.

[3] HERSTEIN, I. N., A commutativity theorem, *J. Algebra* **38** (1976), 112—118. *MR* **53** #549.

[4] HERSTEIN, I. N., *Rings with involution,* Chicago lectures in mathematics, University of Chicago Press, Chicago, Ill.—London, 1976. *MR* **56** #406.

[5] KLEIN, A. A., NADA, I. and BELL, H., Some commutativity results for rings, *Bull. Austral. Math. Soc.* **22** (1980), 285—289. *MR* **82b:** 16038.

(Received January 30, 1986)

DEPARTMENT OF MATHEMATICAL SCIENCES
UNIVERSITY OF PETROLEUM AND MINERALS
DHAHRAN 31261
SAUDI ARABIA

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CALIFORNIA
SANTA BARBARA, CA 93106
U.S.A.

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 37—42

ON A PROBLEM OF BASES FOR THE REGULAR EXTENSION OF VARIETIES OF ALGEBRAS

EWA GRACZYŃSKA

Abstract

For a variety V of algebras of a given type τ, without nullary operations, $R(V)$ denotes the set of all regular identities satisfied in V. The variety $R(V)^*$ of type τ, determined by $R(V)$ is called the regular extension of V. Our aim is to prove a generalization of the result of H. Lakser, R. Padmanabhan and C. R. Platt, announced in [6] and some results of I. I. Mel'nik [7] and E. Graczyńska [2]. Our Theorem 2 gives a partial answer to the probem of determining a base for the regular extension of varieties of algebras, posed in [4].

§ 1. Preliminaries

We deal with universal algebras of a given type $\tau\colon T\to N$, where N denotes the set of all positive integers. Our nomenclature is basically that of [5]. For a given variety of type τ, $E(V)$ denotes the set of all identities of type τ, satisfied in V. If Σ is a set of identities of type τ, then Σ^* denotes the variety of type τ defined by Σ. $E(\Sigma)$ denotes the set of all consequences of Σ. If p is a term of a given type, then $\operatorname{Var}(p)$ denotes the set of all variables occurring in p. An identity $p=q$ is called regular, if $\operatorname{Var}(p)=\operatorname{Var}(q)$ (see [9], [10]). The set of all regular identities of type τ is denoted by $R(\tau)$. For a variety V, $R(V)=E(V)\cap R(\tau)$. If $\underline{A}$ is an algebra of type τ, then $E(\underline{A})$, $R(\underline{A})$ denote the set of all (all regular) identities satisfied in $\underline{A}$, respectively. For a semilattice-ordered system $\mathscr{A}$ of algebras, $S(\mathscr{A})$ denotes the sum of it, according to the definition of Płonka (see [9], [10]). We shall use the name "Płonka sum" instead of "the sum of a semilattice-ordered system of algebras". We refer the reader to our paper [3] for further references on regular identities. For a variety V, $L(V)=(L(V), \cap, \vee)$ denotes the lattice of all subvarieties of V.

Our aim is to prove some results on the regular extension of varieties of algebras. Our Theorem 1 is a generalization of Theorem 2 of H. Lakser, R. Padmanabhan and C. R. Platt, announced in [6] and proved in [1]. Theorem 2 is a generalization of Proposition 2 of I. I. Mel'nik [7] and of Theorem 2 and 3 of [2].

Firstly we recall the definition of Płonka sum:

DEFINITION 1 (Płonka). Let $(\underline{A}_i\colon i\in I)$ be a family of algebras of some similarity type τ, indexed by the elements of a join-semilattice I. Suppose that for each pair $i,j\in I$ with $i\leqq j$ in I a homomorphism $h_{ij}\colon \underline{A}_i\to\underline{A}_j$ is given, so that h_{ii} is the identity map on A_i and $i\leqq j\leqq k$ in I implies $h_{ik}=h_{jk}h_{ij}$. The *Płonka sum* is the

---

1980 *Mathematics Subject Classification* (1985 Revision). Primary 08BXX.

Key words and phrases. Universal algebras, basis, identities, regular identities, varieties of algebras.

algebra $\underline{A}$ of type τ defined on $\bigcup(A_i : i \in I)$ by

$$f\ (a_{i_1}, \ldots, a_{i_n}) = f_t^k(h_{i_1k}(a_{i_1}), \ldots, h_{i_nk}(a_{i_n}))$$

whenever $t \in T, f_t$ is an n-ary operation symbol of type τ, $a_{i_j} \in A_{i_j}$ for each $j = 1, 2, \ldots, n$ and $k = \text{l.u.b.}(i_1, i_2, \ldots, i_n)$; f_t^k denotes the realization of f_t in $\underline{A}_k$, $f_t^{\underline{A}}$ denotes the realization of f_t in $\underline{A}$.

The notion of generalized P-function was investigated in [8]:

Definition 2 (Mitschke). Let $\underline{A} = (A, (f_t : t \in T))$ be an algebra of type $\tau = (n_t : t \in T)$ and f a binary function in A. Then f is called a *generalized P-function* of $\underline{A}$ iff the following equations are satisfied by f, for any $t \in T$:

(1) $$f(f(x, y), z) = f(x, f(y, z)),$$

(2) $$f(f(x, y), f(x, y)) = f(x, y),$$

(3) $$f(x, f(y, z)) = f(x, f(z, y)),$$

(4) $$f(f_t(x_1, \ldots, x_{n_t}), y) = f_t(f(x_1, y), \ldots, f(x_{n_t}, y)),$$

(5) $$f(x, f_t(x_1, \ldots, x_{n_t})) = f(x, f_t(f(x, x_1), \ldots, f(x, x_{n_t}))),$$

(6) $$f(f_t(x_1, \ldots, x_{n_t}), x_k) = f(f_t(x_1, \ldots, x_{n_t}), f_t(x_1, \ldots, x_{n_t})) \quad \text{for} \quad 1 \leqq k \leqq n_t,$$

(7) $$f(x, f_t(x, \ldots, x)) = f(x, x),$$

(8) $$f(f_t(x_1, \ldots, x_{n_t}), f_t(x_1, \ldots, x_{n_t})) = f_t(x_1, \ldots, x_{n_t}) \quad \text{for} \quad n_t \geqq 2.$$

In the sequel we shall write $x \circ y$ instead of $f(x, y)$.

In our considerations we use the following result of [8]:

Theorem (Mitschke). *Let $\underline{A} = (A, (f_t : t \in T))$ be an algebra of type $\tau = (n_t : t \in T)$ and $\circ$ be a generalized P-function of $\underline{A}$. Then $\underline{A}$ is a Płonka sum of subalgebras of $\underline{A}$ satisfying the equality $x \circ y = x \circ x$.*

In the sequel we shall deal with generalized P-functions generated by a binary term $x \circ y$ of type τ. Consequently, $x \circ y$ will denote a binary term of a given type and a generalized P-function.

The notion of generalized P-function is a generalization of the concept of P-function, defined by J. Płonka in [9]. The theorem above is a generalization of Theorem II of [10]. For several examples of algebras with generalized P-function, which is not a P-function, see [8]. In [1], we called a variety V to be strongly nonregular if there exists a binary polynomial symbol p, such that x and y are variables occurring in p and the identity $p(x, y) = x$ holds in V. Obviously, if V is a strongly nonregular variety, then the polynomial symbol $p(x, y)$ generates a generalized P-function in any algebra of V.

§ 2. The lattice $L(R(V)^*)$

Let V be a variety of algebras of type τ: $T\to N$. Assume that there exists a binary term $x\circ y$ of type τ, such that x, y are different variables and $y\in\operatorname{Var}(x\circ y)$ and $x\circ y$ generates a generalized P-function in all algebras $\underline{A}$ of V (i.e. identities (1)—(8) are satisfied in V, for $t\in T$). In this case we shall say, that V is a variety with generalized P-function generated by the term $x\circ y$.

Lemma 1. *Let V be a variety of type τ with generalized P-function generated by the term $x\circ y$, such that the identity $x\circ y=x\circ x$ is satisfied in V. Then the regular extension of V (i.e. the variety defined by $R(V)$) consists of Płonka sums of algebras of V.*

Proof. Obviously, all algebras, which are Płonka sums of algebras of V, belongs to $R(V)^*$ (by Theorem I of [9]). Assume now, that $\underline{A}$ is an algebra of type τ, such that $R(V)\subset E(\underline{A})$. Notice, that identities (1)—(8) are all regular, thus A satisfies all of them. This means that the term $x\circ y$ generates a generalized P-function in $\underline{A}$. By the theorem of Mitschke, the algebra $\underline{A}$ is a Płonka sum of its subalgebras $\underline{A}_i$ satisfying the identity $x\circ y=x\circ x$, $i\in I$. Thus, $\{x\circ y=x\circ x\}\cup R(V)\subset E(\underline{A}_i)$, for all subalgebras $\underline{A}_i$ of the suitable decomposition of $\underline{A}$. By our Corollary from [1], we deduce that all components $\underline{A}_i$ satisfy all identities of $E(V)$, i.e. $\underline{A}$ is a Płonka sum of algebras of V.

The next lemma is a generalization of our Theorem 2 of [1]:

Lemma 2. *Let V be a variety of type τ with generalized P-function generated by the term $x\circ y$, such that the identity $x\circ y=x\circ x$ is satisfied in V. Then the mapping $K\to R(K)^*$ is a monomorphism of the lattice $L(V)$ into the lattice $L(R(V)^*)$.*

Proof. Let $R(K_1)^*=R(K_2)^*$, for $K_1, K_2\in L(V)$. Thus $R(K_1)=R(K_2)$. But $K_i=(R(K_i)\cup\{x\circ y=x\circ x\})^*$, for $i=1,2$ (by Corollary of [1]). Thus $E(K_1)=$ $=E(K_2)$, i.e. $K_1=K_2$. This proves, that the mapping is one-to-one. Obviously, $R(K_1\vee K_2)=R(K_1)\cap R(K_2)$, i.e. $R(K_1\vee K_2)^*=R(K_1)^*\vee R(K_2)^*$. Immediately, we obtain the inclusion $R(K_1\cap K_2)^*\subset R(K_1)^*\cap R(K_2)^*$. Let now $\underline{A}\in R(K_1)^*\cap R(K_2)^*$. By Lemma 1, we obtain, that $\underline{A}$ is a Płonka sum of algebras $\underline{A}_i$, $i\in I$ from the class K_1 and $\underline{A}$ is a Płonka sum of algebras B_j, $j\in J$ from K_2, where I and J are join-semilattices. Using Lemma 1 of [1], we shall prove that this two decompositions of an algebra $\underline{A}$ are equal (i.e. the partitions $\{A_i\colon i\in I\}$ and $\{B_j\colon j\in J\}$ are the same). Let $i\in I$ and $a\in A_i$. Then $a\in B_j$ for some $j\in J$. If b is any element of B_j, then there exists $k\in I$ such that $b\in A_k$. In our case the identity $x\circ y=x\circ x$ is satisfied in K_1 and in K_2. We have two possibilities: 1° $\operatorname{Var}(x\circ y)=\{x, y\}$, 2° $\operatorname{Var}(x\circ y)=$ $=\{y\}$ (by our assumption that $y\in\operatorname{Var}(x\circ y)$). By Lemma 1 of [1], in case 1°, we obtain $a\circ a=a\circ b\in A_m$ and $b\circ b=b\circ a\in A_m$ with $m=\text{l.u.b.}(i,k)$ in I. But $a\circ a\in A_i$ and $b\circ b\in A_k$. Thus $i=k$ and finally $b\in A_i$, which gives $B_j\subset A_i$. By symmetry, $A_i\subset B_j$, i.e. $A_i=B_j$. In case 2°, we obtain $a\circ a=a\circ b\in A_k$ (by Lemma 1 of [1]), but $a\circ a\in A_i$, so $i=k$. As before, we get $A_i=B_j$. So we have proved that $\underline{A}_i$ belongs to the variety $K_1\cap K_2$ for $i\in I$. Finally, by Theorem I of [9] we conclude that $\underline{A}\in R(K_1\cap K_2)^*$. ∎

Theorem 1. *Let V be a variety of type τ with generalized P-function generated by the term $x\circ y$, such that the identity $x\circ y=x\circ x$ is satisfied in V. Then the lattice*

$L(R(V)^*)$ is isomorphic to the direct product of the lattice $L(V)$ and a two element-chain.

PROOF. Let us agree to use the symbol 2 for a two-element lattice $(\{0, 1\}, \wedge, \vee)$ with greatest element 1. Define $g\colon L(V)\times 2\to L(R(V)^*)$ by $g(\langle K, 0\rangle)=K$ and $g(\langle K, 1\rangle)=R(K)^*$ for $K\in L(V)$. By Lemma 2, g is one-to-one. To prove, that g is a lattice homomorphism, we use Theorem 1 of [1] and the following equalities, for $K_1, K_2\in L(V)$: $R(K_1\vee K_2)^*=K_1\vee R(K_2)^*$, $K_1\cap K_2=K_1\cap R(K_2)^*$. To complete the proof it suffices to show that g is onto $L(R(V)^*)$. Clearly, every element $K\in L(V)$ is the image by g of $\langle K, 0\rangle$. Now assume, that K' is an element of $L(R(V)^*)$, but is not an element of $L(V)$. Thus K' is generated by some algebra $\underline{A}'$ of K', i.e. $K'=\mathrm{HSP}(\underline{A}')$. But $K'\subset R(V)^*$, thus $\underline{A}'\in R(V)^*$ and by our Lemma 1 it follows that $\underline{A}'$ is a Płonka sum of algebras $\underline{A}_i$, $i\in I$ from V. But $\underline{A}'$ is not in V, thus $\operatorname{card}(I)>1$ and $E(K')=E(\underline{A}')=R(\underline{A}')=\bigcap(R(\underline{A}_i\colon i\in I))$, by Theorem 1 of [9]. But $\bigcap(R(\underline{A}_i)\colon i\in I)=R(\prod(\underline{A}_i\colon i\in I))$ and the algebra $\underline{A}=\prod(\underline{A}\colon i\in I)$ is in V. Take the variety K generated by $\underline{A}$, i.e. $K=\mathrm{HSP}(\underline{A})$. We obtain that $K\in L(V)$ and $E(K)=E(\underline{A})$, i.e. $R(K)=R(\underline{A})=R(\underline{A}')=E(K')$. This implies $R(K)^*=K'$, hence $K'=g(\langle K, 1\rangle)$. ∎

§ 3. Bases for the regular extension of a variety

Several interesting varieties of algebras are defined by a set of regular identities and sometimes they constitute the regular extension of another variety. Thus the problem arises to determine a base for the variety $R(V)^*$ if a base for V is given. This is solved in cases (1)—(4) of the Lemma of [10], by I. I. Mel'nik, J. Płonka and the author. In particular, in Cases (1)—(4) of [10], one obtains that for varieties of finite type (i.e. when T is finite) the regular extension of a variety V is finitely based iff V is finitely based. However, in the remaining case (i.e. in case (5) of [10]) this problem is more complicated. In [4] we presented two varieties of algebras: one of type (2), the second one of type (2, 1), both defined by the single identity $x\cdot y=x\cdot x$ and we have observed that the regular extension of the first one is finitely based but the regular extension of the second is not finitely based. Thus we expected that conditions on varieties that guarantee that their regular extensions are finitely based must involve the similarity type. The following theorem presents such conditions:

THEOREM 2. *Let V be a variety of type τ, with generalized P-function generated by term $x\circ y$, such that the identity $x\circ y=x\circ x$ is satisfied in V. Assume, that V is defined by a set Σ of regular identities and the identity $x\circ y=x\circ x$. Then the regular extension of V is defined by $\Sigma\cup\{(1)\text{—}(8), \text{for } t\in T\}$ (from Definition 2 of the generalized P-function for $f(x, y)$ being $x\circ y$).*

PROOF. Denote by Q the set of identities $\Sigma\cup\{(1)\text{—}(8)\colon t\in T\}$. Our aim is to show that Q is a base for the variety $R(V)^*$ of type τ. By definition, all identities of Q are regular and they are valid in V. Thus the inclusion $Q\subset R(V)$ holds, i.e. $R(V)^*\subset Q^*$. To prove the converse inclusion, assume that $\underline{A}$ is an algebra of type τ and $\underline{A}$ is in Q^*, i.e. $\underline{A}$ satisfies all identities of Q. This means that the term $x\circ y$ generates a generalized P-function in $\underline{A}$. By A. Mitschke's theorem we conclude that $\underline{A}$ is a Płonka sum of algebras $\underline{A}_i$, $i \in I$, which are subalgebras of $\underline{A}$ and satisfy the identity $x \circ y = x \circ x$. Thus all algebras $\underline{A}_i$, $i \in I$ satisfy all identities of Q and the identity $x \circ y = x \circ x$. But Σ is included in Q and by our assumption, we get that all algebras $\underline{A}_i$, $i \in I$ belong to the variety V. Thus $\underline{A}$ belongs to $R(V)^*$, by Theorem I of [9]. Finally we have $R(V)^* = Q^*$. ∎

REMARK. The assumption that the term $x \circ y$ generates a generalized P-function in the algebras of V, is essential. One can easily see, that our variety K_0 of type (2, 1) defined by the single axiom $x \cdot y = x \cdot x$, is not a variety with a generalized P-function generated by the term $x \cdot y$, because its regular extension does not constitute the variety of Płonka sums of algebras of K_0 (see Theorem 1 of [4]).

COROLLARY. *Let V be a variety of type τ, with generalized P-function generated by a term $x \circ y$, such that the identity $x \circ y = x \circ x$ is satisfied in V. Assume that the type τ is finite (i.e. the set T is finite). Then V is finitely based if and only if $R(V)^*$ is finitely based. Analogously, V has the finite basis property if and only if $R(V)^*$ has this property (cf. with the Corollary and with the proof of Remark 4 of* [2]*).*

PROOF. The first statement of the Corollary follows from Theorem 2 and the fact that $R(V) \cup \{x \circ y = x \circ x\}$ is a base for V, so if V has a finite base then there exists a finite subset Σ of $R(V)$ such that $\Sigma \cup \{x \circ y = x \circ x\}$ is a base for V. A method, how to find such a subset Σ of $R(V)$ was shown in [2], in general case. If the term $x \circ y$ is a generalized P-function, the following method was proposed by A. Mitschke: assume that a finite base of V is of the form $\Sigma_1 \cup \Sigma_2$, where $\Sigma_1 \subset R(V)$ and $\Sigma_2 \subset E(V) - R(V)$. Let Σ_3 be the set of identities constituted in the following way: if

$$f(x_1, \dots, x_n, y_1, \dots, y_m) = g(x_1, \dots, x_n, z_1, \dots, z_k) \in \Sigma_2$$

where $x_1, \dots, x_n$ are all the variables which occur both in f and in g, then

$$f(x_1, \dots, y_m) \circ z_1 \circ \dots \circ z_m = g(x_1, \dots, z_k) \circ y_1 \circ \dots \circ y_k \in \Sigma_3.$$

Now let $\Sigma = \Sigma_1 \cup \Sigma_3 \cup \{(1)\text{—}(8)\colon t \in T\}$. Then using the properties of the term $x \circ y$, as in the proof of Satz 2.7 of [8], one can prove that Σ is a base for the variety $R(V)^*$.

Recall that a variety V is said to have the finite basis property, if the set of identities of any finite member of V is finitely based. Assume, that V has the finite basis property. Take a finite algebra $\underline{A}$ of $R(V)^*$. By Lemma 1, $\underline{A}$ is a Płonka sum of algebras $\underline{A}_i$, $i \in I$. Let $\underline{B} = \prod (\underline{A}_i \colon i \in I)$. Obviously, $\underline{B}$ is a finite algebra of V. If $\operatorname{card}(I) = 1$, then $\underline{A} \in V$ and $\underline{A}$ is finitely based by assumption. Otherwise, $E(\underline{A}) = R(\underline{B})$, where $R(\underline{B})$ is the set of all regular identities satisfied in $\underline{B}$, by Theorem 1 of [9]. Let K be the variety of type τ generated by $\underline{B}$. Then $E(\underline{A}) = R(\underline{B}) = R(\tau) \cap E(K) = R(K)$ and by Theorem 2 we conclude that $E(\underline{A})$ is finitely based. ∎

ACKNOWLEDGEMENTS. This work was done while the author was visiting the Technische Hochschule Darmstadt, under the auspices of the Alexander von Humboldt Foundation. The author is very grateful to dr. Aleit Mitschke for her valuable comments during writing this paper.

REFERENCES

[1] DUDEK, J. and GRACZYŃSKA, E., The lattice of varieties of algebras, *Bull. Acad. Polon. Sci. Sér. Sci. Math.* **79** (1981), 337—340. *MR* **83a:** 08013.

[2] GRACZYŃSKA, E., On regular identities, *Algebra Universalis* **17** (1983), 369—375. *MR* **86a:** 08006.

[3] GRACZYŃSKA, E., Regular identities (submitted).

[4] GRACZYŃSKA, E., KELLY, D. and WINKLER, P., On the regular extension of varieties of algebras, *Algebra Universalis* **23** (1986), 77—84.

[5] GRÄTZER, G., *Universal algebra*, 2nd ed., Springer-Verlag, New York—Heidelberg, 1979. *MR* **80g:** 08001.

[6] LAKSER, H., PADMANABHAN, R. and PLATT, C. R., Algebras defined by regular identities, *Notices Amer. Math. Soc.* **17** (1970), p. 945 (abstract).

[7] MEL'NIK, I. I., Normal closure of perfect varieties of universal algebras, *Ordered sets and lattices*, No. 1, Izdat. Saratov. Univ., Saratov, 1971, 56—65 (in Russian). *MR* **52** # 3011.

[8] MITSCHKE, A., Summen von gerichteten Systemen von Algebren, *Mitt. Math. Sem. Giessen*, Heft **124**, 1977, 1—56. (Dissertation des Fachbereich Mathematik der Technischen Hochschule Darmstadt, April 1976.) *MR* **56** # 2899.

[9] PŁONKA, J., On a method of construction of abstract algebras, *Fund. Math.* **61** (1967), 183—189. *MR* **37** # 1294.

[10] PŁONKA, J., On equational classes of abstract algebras defined by regular equations, *Fund. Math.* **64** (1969), 241—247. *MR* **39** # 5450.

(Received February 5, 1986)

TECHNISCHE HOCHSCHULE DARMSTADT
FACHBEREICH MATHEMATIK, A. G. 1
SCHLOSSGARTENSTRASSE 7
D—6100 DARMSTADT
FEDERAL REPUBLIC GERMANY

Current address:

INSTYTUT MATEMATYCZNY
POLSKIEJ AKADEMII NAUK
UL. KOPERNIKA 18
PL—51-617 WROCŁAW
POLAND

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 43—45

ON SEMINORM SEPARABILITY FOR VECTOR-VALUED FUNCTION SPACES

LIAQAT ALI KHAN

Abstract

Let X be a completely regular Hausdorff space and E a Hausdorff locally convex space, and let σ_0, σ, and u denote the countable-open, σ-compact-open, and uniform topologies, respectively on $C_b(X, E)$. Let $C_{rc}(X, E)$ be the subspace of $C_b(X, E)$ consisting of those functions f for which $f(X)$ is relatively compact. Then (1) $(C_{rc}(X, E), \sigma_0)$ or $(C_{rc}(X, E), \sigma)$ is seminorm separable iff the closure in βX of every σ-compact subset of X is metrizable in βX and E is seminorm separable; (2) $(C_b(X, E), \sigma_0)$ is seminorm separable iff $(C_b(X, E), \sigma)$ is so; (3) $(C_b(X, E), u)$ is seminorm separable iff X is a compact metric space and E is seminorm separable.

Let $C_b(X, E)$ (resp. $C_{rc}(X, E)$) denote the vector space of all continuous E-valued functions f on X such that $f(X)$ is bounded (resp. relatively compact). When E is the real or complex field, we write $C_b(X)$ for $C_b(X, E)$. The *countable-open topology* σ_0 *(resp. σ-compact-open topology σ)* [2, 3, 6] on $C_b(X, E)$ is generated by the family $\{\|\cdot\|_{A,p}\}$ of seminorms, where A varies over all countable (resp. σ-compact) subsets of X, p varies over a family P of continuous seminorms generating the topology of E, and $\|f\|_{A,p}=\sup\{p(f(x))\colon x\in A\}$, $f\in C_b(X, E)$. The *uniform topology* u on $C_b(X, E)$ is given by the seminorms $\|\cdot\|_p=\|\cdot\|_{X,p}$, $p\in P$. When these topologies are considered on $C_b(X)$, we shall omit the suffix p. It is easily seen that $\sigma_0\subseteq\sigma\subseteq u$. We shall denote by $C_b(X)\otimes E$ the vector space spanned by the set of all functions of the form $g\otimes a$, where $g\in C_b(X)$, $a\in E$, and $(g\otimes a)(x)=g(x)a$ $(x\in X)$. The Stone—Čech compactification of X is denoted by βX.

A locally convex space L is called *seminorm separable* if, for every continuous seminorm q on L, (L, q) is separable. The following two results, due to M. and S. Krein [8] and Gulick and Schmets [3], are stated for reference purpose.

THEOREM 1. [8] $(C_b(X), \|\cdot\|)$ *is separable iff X is a compact metric space.*

THEOREM 2. [3] *The following statements are equivalent.*

(a) $(C_b(X), \sigma)$ *is seminorm separable.*
(b) $(C_b(X), \sigma_0)$ *is seminorm separable.*
(c) *The closure in βX of each σ-compact subset of X is metrizable in βX.*

Our purpose is to extend these results to vector-valued function spaces. We shall require the following known result (see [7, p. 286] or [4, Lemma 2.2]).

THEOREM 3. $C_b(X)\otimes E$ *is u-dense, hence σ_0- and σ-dense, in* $C_{rc}(X, E)$.

1980 *Mathematics Subject Classification.* Primary 46E10; Secondary 46E40, 46A05.
Key words and phrases. Locally convex spaces, vector-valued continuous functions, seminorm separable spaces, countable-open and σ-compact-open topologies, Stone—Čech compactification.

PROOF. If Y is any compact Hausdorff space, then by [1, Ch. III. 1, Prop. 1 and Lemma 2] $C_b(Y)\otimes E$ is u-dense in $C_b(Y, E)$. Since each f in $C_b(X)$ or $C_{rc}(X, E)$ has a continuous extension to all of βX, it follows that $C_b(X)$ and $C_{rc}(X, E)$ are linearly isomorphic to $C_b(\beta X)$ and $C_b(\beta X, E)$, respectively. Consequently, $C_b(X)\otimes E$ is u-dense in $C_{rc}(X, E)$.

THEOREM 4. *The following statements are equivalent.*

(a) $(C_{rc}(X, E), \sigma)$ *is seminorm separable.*

(b) $(C_{rc}(X, E), \sigma_0)$ *is seminorm separable.*

(c) *The closure in* βX *of every* σ*-compact subset of* X *is metrizable in* βX *and* E *is seminorm separable.*

PROOF. (a)⇒(b) This follows easily from the fact that $\sigma_0\subseteq\sigma$.

(b)⇒(c) Suppose that $(C_{rc}(X, E), \sigma_0)$ is seminorm separable, and let A be a σ-compact subset of X. Choose $\varphi\in E'$ and $c\in E$ with $\varphi(c)=1$. There exists some $p\in P$ such that $|\varphi(a)|\leqq p(a)$ for all $a\in E$. Let $\{f_n\}$ be a dense subset of $(C_{rc}(X, E), \|\cdot\|_{A,p})$. Then $\{\varphi\circ f_n\}$ is dense in $(C_b(X), \|\cdot\|_A)$ as follows. Let $g\in C_b(X)$ and $\varepsilon>0$. Choose an integer N such that $\|f_N-g\otimes c\|_{A,p}<\varepsilon$. Then, for any $x\in A$,

$$|\varphi(f_N(x))-g(x)| = |\varphi(f_N(x)-g(x))| \leqq p(f_N(x)-g(x)c) < \varepsilon.$$

Hence $(C_b(X), \sigma)$ is seminorm separable, and so the first part of (c) follows from Theorem 2. Next, let $p_1\in P$. For any fixed $z\in X$, let $\{g_n\}$ be a dense subset of $(C_{rc}(X, E), \|\cdot\|_{\{z\},p_1})$. Then $\{g_n(z)\}$ is dense in (E, p_1), and this proves the second part of (c).

(c)⇒(a) Let A be a σ-compact subset of X, and let $p\in P$. By Theorem 2, $(C_b(X), \|\cdot\|_A)$ is separable and so it has a countable dense subset $\{h_m\}$ (say). Let $\{a_n\}$ be a dense subset of (E, p). Let H be the countable subspace generated by $\{h_m\otimes a_n\colon m, n=1, 2, \ldots\}$ over rationals. In view of Theorem 3, it suffices to show that H is dense in $(C_b(X)\otimes E, \|\cdot\|_{A,p})$. Let $g=\sum_{i=1}^{k} g_i\otimes b_i$ $(g_i\in C_b(X),\ b_i\in E)$ be in $C_b(X)\otimes E$ and $0<\varepsilon<1$. Let $r=\max\{\|g_i\|, p(b_i)\colon 1\leqq i\leqq k\}$. For each $i=1, \ldots, k$, choose $h_{m_i}\in\{h_m\}$ and $a_{n_i}\in\{a_n\}$ such that

$$\|h_{m_i}-g_i\|_A < \varepsilon/2k(r+1) \quad\text{and}\quad p(a_{n_i}-b_i) < \varepsilon/2k(r+1).$$

Note that, for each $x\in A$ and $1\leqq i\leqq k$, $|h_{m_i}(x)|<r+1$. Let $h=\sum_{i=1}^{k} h_{m_i}\otimes a_{n_i}$. Then $h\in H$ and, for any $x\in A$,

$$p(h(x)-g(x)) \leqq \sum_{i=1}^{k} |h_{m_i}(x)|\,p(a_{n_i}-b_i)+\sum_{i=1}^{k} |h_{m_i}(x)-g_i(x)|\,p(b_i) < \varepsilon.$$

This completes the proof.

REMARK. If X is a normal or a locally compact space, then for any σ-compact $A\subseteq X$, the closure of A in βX coincides with that in X [3, Theorem 9]. Under this assumption, the condition (c) of Theorem 4 may be replaced by:

(c′) *The closure in X of each σ-compact subset of X is compact and metrizable and E is seminorm separable.*

We do not know whether $C_b(X)\otimes E$ is σ-dense in $C_b(X, E)$. If this is so, then indeed Theorem 4 holds with $C_{rc}(X, E)$ replaced by $C_b(X, E)$. However, analogous to [3, Theorem 4], we obtain

THEOREM 5. *$(C_b(X, E), \sigma)$ is seminorm separable iff $(C_b(X, E), \sigma_0)$ is so.*

PROOF. Suppose $C_b(X, E), \sigma_0)$ is seminorm separable but $(C_b(X, E), \sigma)$ is not. Then there exist some $p\in P$ and σ-compact set $A\subseteq X$ such that $(C_b(X, E), \|\cdot\|_{A,p})$ is not separable. Consequently, for any sequence $\{f_n\}\subseteq C_b(X, E)$, there exist $\varepsilon>0$, $f\in C_b(X, E)$ and a sequence $\{x_n\}\subseteq A$ such that $p(f_n(x_n)-f(x_n))\geqq\varepsilon$ for all n. Then, if $D=\{x_n\}$, $\|f_n-f\|_{D,p}\geqq\varepsilon$ for all n. This implies that $(C_b(X, E), \sigma_0)$ is not seminorm separable.

Finally, by Theorem 1 and the argument used in the proof of Theorem 4, we can easily establish the following (cf. [5, Theorem 3]).

THEOREM 6. *$(C_b(X, E), u)$ is seminorm separable iff X is a compact metric space and E is seminorm separable.*

REFERENCES

[1] BOURBAKI, N., Éléments de mathématique. Fasc. XIII. Livre VI: *Intégration.* Chapitres 1, 2, 3 et 4, Deuxième édition rev. et augmentée, Actualités Sci. Indust., no. 1175, Hermann, Paris, 1965. *MR* **36** #2763.

[2] GULICK, D., The σ-compact-open topology and its relatives, *Math. Scand.* **30** (1972), 159—176. *MR* **48** #9366.

[3] GULICK, D. and SCHMETS, J., Separability and semi-norm separability for spaces of bounded continuous functions, *Bull. Soc. Roy. Sci. Liège* **41** (1972), 254—260. *MR* **49** # 5795.

[4] KATSARAS, A. K., Continuous linear functionals on spaces of vector-valued functions, *Bull. Soc. Math. Grèce (N. S.)* **15** (1974), 13—19. *MR* **58** #30159.

[5] KHAN, L.A., Separability in the uniform topology, *Studia Sci. Math. Hungar.* **20** (1985), 407—409.

[6] KHAN, L. A., The countable-open topology in the locally convex setting, *Kobe J. Math.* **3** (1986), 47—50. *MR* **87k:** 46052.

[7] KLEE, V., Leray—Schauder theory without local convexity, *Math. Ann.* **141** (1960), 286—296. *MR* **24** #A1004.

[8] KREIN, M. and KREIN, S., On an inner characteristic of the set of all continuous functions defined on a bicompact Hausdorff space, *C. R. (Doklady) Acad. Sci. URSS (N. S.)* **27** (1940), 427—430. *MR* **2**—222.

(Received February 10, 1986)

DEPARTMENT OF MATHEMATICS
FEDERAL GOVERNMENT COLLEGE
H—8
ISLAMABAD
PAKISTAN

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 47—65

DETERMINING SMOOTHNESS BY BLOCK DATA

Z. DITZIAN

1. Introduction

The smoothness of $f(x)$ in $L_p(D)$ is measured by

$$\omega_r(f, t)_p = \sup_{0<h\leq t} \|\Delta_h^r f\|_{L_p(D)} \tag{1.1}$$

where

$$\Delta_h^r f(x) = 0 \quad \text{for} \quad [x-rh/2,\ x+rh/2] \not\subset D$$

and

$$\Delta_h^r f(x) = \sum_{k=0}^{r} \binom{r}{k} (-1)^k f(x-kh+rh/2) \quad \text{for} \quad [x-rh/2,\ x+rh/2] \subset D. \tag{1.2}$$

In the definition of $\omega_r(f, t)_p$ given above, information on all h and on $|\Delta_h^r f(x)|$ for almost all x is used. The problem of determining $\omega_r(f, t)_p$ when $\|\Delta_{h_n}^r f\|$ is known for a sequence h_n satisfying $h_n = o(1)$ as $n\to\infty$ and $h_n/h_{n+1} = O(1)$ as $n\to\infty$ was investigated by R. DeVore [5, p. 258], G. Freud [11], J. Boman [1], V. Totik [13] and the author [6] and [9]. It is well-known that the condition $h_n/h_{n+1} = O(1)$ cannot be dropped and examples were given for which $\|\Delta_{h_n}^r f\|_p = O(h_n^\alpha)$, $\alpha < r$ for some fast decreasing sequence h_n but $\|\Delta_h^r f\|_p \neq O(h^\alpha)$. (The condition $h_n/h_{n+1} = O(1)$ can be omitted for $\|\Delta_{h_n}^r f\|_p = O(h_n^r)$.)

For the space $C(D)$ investigations were carried by Z. Ciesielski [2], [3] and the author [8] to show how data on a given sequence of points can be sufficient to determine $\omega_r(f, t)_\infty$ in many cases.

For a function in L_p the information should be on averages on some given blocks rather than on discrete data. The main results and an outline of the various stages of the proof will be given in Section 2.

2. The main result

For the sequences $\xi_n \in R$, and $h_n \in R$ where $h_n = o(1)$, $n\to\infty$, we define the block averages by

$$a_k(n) = h_n^{-1} \int_0^{h_n} f(\xi_n + kh_n + u)\, du. \tag{2.1}$$

Supported by NSERC grant A4816 of Canada.

1980 *Mathematics Subject Classification* (1985 Revision). Primary 26A16; Secondary 41A25, 41A27.

Key words and phrases. Moduli of smoothness. spline approximation block data. inverse theorem.

We emphasize that ξ_n is given as an added degree of freedom, not an added condition, and all theorems here are valid if $\xi_n=0$. We define $\Delta^r a_k(n)$ by

$$\Delta^r a_k(n) = \sum_{l=0}^{r} \binom{r}{l} (-1)^l a_{k+r-l}(n) \tag{2.2}$$

(the usual forward r-difference). For a finite or infinite interval D and for $1 \leqq p < \infty$ we define

$$\|\Delta^r a_k(n)\|_{l_p(D)} = \Big\{ \sum_{k \in \mathcal{K}(n)} |\Delta^r a_k(n)|^p \Big\}^{1/p} \tag{2.3}$$

where $\mathcal{K}(n) = \{k; [\xi_n + kh_n, \xi_n + (k+r+1)h_n] \subset D\}$. For $p=\infty$ we have

$$\|\Delta^r a_k(n)\|_{l_\infty(D)} = \sup_{k \in \mathcal{K}(n)} |\Delta^r a_k(n)|. \tag{2.4}$$

A function φ belongs to class Φ, $\varphi \in \Phi$, if $\varphi(h)$ is increasing in R_+, $\lim_{h \to 0+} \varphi(h) = 0$ and for some $0 < \eta < 1$ there exists a constant $L = L(\eta)$ such that $(\varphi(h)/\varphi(Lh)) \leqq \eta$ for all $h > 0$. This class Φ of functions will be used in this paper to measure smoothness instead of just dealing with h^α. It is obvious that h^α for $\alpha > 0$ is of class Φ. Many other functions, for instance $\varphi(h) = h^\alpha (\log 1/h)^\beta$ where $\alpha > 0$, are of class Φ.

With these notations we can state the main result of the paper.

THEOREM 2.1. *Suppose D is a finite or infinite interval, $f \in L_p(D)$, $1 \leqq p \leqq \infty$, $\{h_n\}$ is a sequence of real numbers satisfying $h_n = o(1)$, $n \to \infty$, and $1 \leqq h_n/h_{n+1} \leqq M$, and $\varphi(h)$ is of class Φ. Then*

$$\|\Delta^r a_k(n)\|_{l_p(D)} = O\big(\varphi(h_n) h_n^{-1/p}\big) \Leftrightarrow \omega_r(f, t)_p = O\big(\varphi(t)\big) \tag{2.5}$$

and

$$\|\Delta^r a_k(n)\|_{l_p(D)} \sim \varphi(h_n) h_n^{-1/p} \Leftrightarrow \omega_r(f, t)_p \sim \varphi(t) \tag{2.6}$$

where $a_n \sim b_n$ if $A^{-1} a_n < b_n < A a_n$.

THEOREM 2.2. *For $p=1$ the result of Theorem 2.1 is valid if we assume in addition either*

(a) $\|\Delta_h f\|_{L_1(D)} \leqq Mh^\delta$ *for some* $\delta > 0$ $(f \in \mathrm{Lip}\, \delta)$;

or

(b) $\xi_n = \xi$ *and* $h_n = m(n) h_{n+1}$ *where $m(n)$ is an integer satisfying* $m(n) \leqq M$.

As a corollary we obtain

COROLLARY 2.3. *Suppose $f \in L_p(D)$, $1 \leqq p \leqq \infty$, and if $p=1$, either* (a) *or* (b) *of Theorem 2.2 is satisfied. Then for $\alpha \leqq r$*

$$\|\Delta^r a_k(n)\|_{l_p(D)} = O(h_n^{\alpha - 1/p}) \Leftrightarrow \omega_r(f, t)_p = O(t^\alpha) \tag{2.7}$$

and

$$\|\Delta^r a_k(n)\|_{l_p(D)} \sim h_n^{\alpha - 1/p} \Leftrightarrow \omega_r(f, t)_p \sim t^\alpha. \tag{2.8}$$

To prove Theorems 2.1 and 2.2 we will need an approximating operator that will be described in Section 3 together with its properties, the proof of which will be given in Section 4. In Section 5 the main theorem will be proved for $D=R$ under the a priori condition $f\in \mathrm{Lip}_p\,\delta$. In Section 6 the above mentioned a priori condition is dropped for $D=R$ and $1<p\leqq\infty$ or replaced by (b) of Theorem 2.2 for $L_1(R)$. In Section 7 the result is extended to domain D with a finite end-point. Finally, in Section 8 we indicate how the ideas of this paper can be used to extend the results of [8] on determining modulus of smoothness using discrete data in $C(D)$.

3. The approximating operator

For the proof of Theorems 2.1 and 2.2 we will need an approximating operator using spline functions. The B spline $G_m(t)$ supported by $[-m/2, m/2]$ with knots at $\frac{m}{2}+j$ $(j=-m, \ldots, 0)$ is given by (see [12, p. 138])

(3.1)

$$G_m(t) = G*G_{m-1}(t) = \int G_{m-1}(u)G(t-u)\,du, \quad G(t) = G_1(t) = \begin{cases} 1 & \text{for } \quad |t| \leqq \dfrac{1}{2} \\ 0 & \text{otherwise.} \end{cases}$$

We define now the approximating operator $S_n(f)$ by

$$S_n(f) = S_n(m, f;\ t) = \sum_k a_k(n) G_m\left(-\left(k+\frac{1}{2}\right)+(t-\xi_n)/h_n\right) \tag{3.2}$$

where $a_k(n)$ is given by (2.1) and recall that $G_m\left(-\left(k+\frac{1}{2}\right)+(t-\xi_n)/h_n\right)$ is the B spline with knots at $\xi_n+\left(\frac{m+1}{2}+j\right)h_n$ supported by

$$\left[\xi_n+kh_n-\frac{m-1}{2}h_n,\ \xi_n+kh_n+\frac{m+1}{2}h_n\right].$$

The main approximating operator of this paper is

$$S_{n,l}(f) = S_{n,l}(m, f;\ t) = \sum_{j=1}^{r} (-1)^{j+1}\binom{l}{j} S_n^j(m, f;\ t) \tag{3.3}$$

where $S_n^j(m, f; t)\equiv S_n^j(f)$, the iteration of $S_n(f)$, is given by

$$S_n^j(f) = S_n(S_n^{j-1}(f)), \qquad S_n^1(f) = S_n(f). \tag{3.4}$$

We can now state the approximation result in $L_p(R)$.

THEOREM 3.1. *For $f\in L_p(R)$, $1\leqq p\leqq\infty$, $r\leqq m$ and $2l>r-1$ we have*

$$\|S_{n,l}(m, f;\ \cdot\,)-f(\cdot\,)\|_p \leqq K\omega_r(f, h_n)_p \tag{3.5}$$

where K is independent of n and f.

The local approximation result is given by

THEOREM 3.2. *Suppose* $f\in L_p[a,b]$, $r\leqq m$ *and* $2l>r-1$. *Then*

$$\|S_{n,l}(m,f;\ \cdot)-f(\cdot)\|_{L_p[a+Ah_n,\,b-Ah_n]}\leqq K\omega_r(f,h_n)_{L_p[a,b]} \tag{3.6}$$

for some A and K which do not depend on n nor on f.

We chose domain $[a,b]$ rather than $[0,1]$ or R_+ as this theorem will be used in Section 7 with different a and b in an inductive manner. In fact, Theorem 2.1 follows from Theorem 2.2. The proof of these results is attained via several lemmas and will be given in the next section.

4. The rate of approximation

In this section we will prove Theorems 3.1 and 3.2 for which we will need the following lemmas:

LEMMA 4.1. *For a polynomial* $p(t)$ *of degree* $\mu<m$ *and* $S_n(f)$ *given in* (3.2)

$$S_n(p)-p=q$$

where q is a polynomial of degree $\mu-2$ *if* $\mu-2\geqq 0$ *and* $q=0$ *if* $\mu-2<0$.

PROOF. This lemma is not dependent on n (or the particular choice of ξ_n and h_n), that is, for $S(f)=S(m,f,t)$ given by

(4.1)
$$S(f)=S(m,f;\ t)=\sum_k\left\{h^{-1}\int_0^h f(\xi+kh+u)\,du\right\}G_m\left(-\left(k+\frac{1}{2}\right)+(t-\xi)/h\right)$$

we have $q=S(p)-p$ is a polynomial of degree $\mu-2$ if $\mu-2\geqq 0$ and is 0 otherwise.

We first show that q is of degree no bigger than $\mu-1$ if $\mu-1\geqq 0$ and observe that as $S(m,1;t)=1$, $\mu=0$ implies $q\equiv 0$. A well-known result about B splines [2, p. 138] yields

$$\frac{d}{dt}S(m,f;\ t)=h^{-1}S(m-1,\vec{\Delta}_h f;\ t), \tag{4.2}$$

and therefore

$$\left(\frac{d}{dt}\right)^l S(m,f;\ t)=h^{-l}S(m-l,\vec{\Delta}_h^l f;\ t) \tag{4.3}$$

where $\vec{\Delta}_h^l f$ is given by

$$(\vec{\Delta}_h^l f(x))=\vec{\Delta}_h(\vec{\Delta}_h^{l-1}f(x))\quad\text{and}\quad\vec{\Delta}_h f(x)=f(x+h)-f(x). \tag{4.4}$$

It is enough to prove the result now for $p(t)=t^l$, $0\leqq l<m$. Using (4.3) for $f(t)=$ $=p(t)=t^l$, $S(m,p(\cdot);t)$ is, between knots, $t^l+q(t)$ where $q(t)$ is a polynomial of degree at most $l-1$. As $S(m,f,t)$ has $m-2$, therefore $l-1$, continuous

derivatives (and for $p(t)=t^l$ we have shown that $\left(\frac{d}{dt}\right)^l S(m, f, t)=l!$), we have for p satisfying $\deg p(t)\leqq\mu$, $0<\mu<m$,

$$S(m, p;\ t)-p(t) = q(t)$$

where q is a polynomial of degree at most $\mu-1$.

We now examine $p(t)=(t-\xi-(h/2))^\mu$ for $\mu<m$. Obviously, the above yields

$$S(m, p;\ t)-p(t) = q_\mu(t-\xi-(h/2))$$

where $q_\mu(u)$ is a polynomial of degree $\mu-1$. As $G_m(u)=G_m(-u)$, (4.1) implies that $S(m, p, t)$ is odd or even in $(t-\xi-(h/2))$ for μ odd or even, respectively. Therefore, $q_\mu(t-\xi-h/2)$ is an odd or even polynomial when μ is odd or even and consequently its degree is at most $\mu-2$. As every polynomial is a combination of terms of the above form, our lemma follows.

Lemma 4.2. *For polynomial* $p(x)$ *of* $\deg\mu$, $\mu<\min(2r, m)$,

$$S_{n,r}(m, p;\ t) = p(t). \tag{4.5}$$

Proof. Using (3.3) and (3.4), our result is reduced to r iteration of Lemma 4.1.

Lemma 4.3. *For* $f\in L_p(R)$

$$\|S_{n,r}(m, f;\ \cdot)\|_p \leqq (2^r-1)\|f(\cdot)\|_p.$$

Proof. Using (3.3) and (3.4), it is enough to show

$$\|S_{n,1}(m, f;\ \cdot)\|_p = \|S_n(m, f;\ \cdot)\|_p \leqq \|f\|_p.$$

Since for $p=\infty$ we have

$$\|S_n(m, f;\ \cdot)\|_\infty \leqq \sup_k |a_k(n)| \sum G_m\left(-\left(k+\frac{1}{2}\right)+(t-\xi_n)/h_n\right) = \sup_k |a_k(n)| \leqq \|f\|_\infty,$$

and for $p=1$ we have

$$\|S_n(m, f;\ \cdot)\|_1 \leqq \sum_k |a_k(n)|\, h_n \leqq \|f\|_1,$$

the Riesz—Thorin interpolation theorem yields our lemma.

We define $D(\zeta)$ by $[a+\zeta, b-\zeta]$ for $D=[a, b]$ and by $[a+\zeta, \infty)$ for $D=[a, \infty)$.

Lemma 4.4. *For* $f\in L_p(D)$ *where* $D=[a, b]$ *or* $D=[a, \infty)$

$$\|S_{n,r}(m, f;\ \cdot)\|_{L_p(D(Ah_n))} \leqq (2^r-1)\|f\|_{L_p(D)} \tag{4.7}$$

where A *depends only on* m *and* r.

Proof. The functional $S_n(m, f, t)$ depends only on the values of $f(x)$ in the interval $[t-(m+1)h_n/2, t+(m+1)h_n/2]$. (The exact constant $(m+1)/2$ is not important.) Therefore, $S_n^j(m, f; t)$ depends only on the values of $f(x)$ in $[t-j(m+1)h_n/2, t+j(m+1)h_n/2]$ and $S_{n,r}(m, f; t)$ depends on those in

$$[t-r(m+1)h_n/2,\ t+r(m+1)h_n/2].$$

We choose A to be any number bigger than $r(m+1)/2$. We define $f^*\in L_p(R)$ by $f^*=0$ if $x\notin D$ and by $f^*=f$ if $x\in D$, and as $S_{n,r}(m,f^*;t)=S_n(m,f;t)$ for $t\in D(Ah_n)$, we have

$$\|S_{n,r}(m,f;\ \cdot)\|_{L_p(D(Ah_n))} = \|S_{n,r}(m,f^*;\ \cdot)\|_{L_p(D(Ah_n))} \leqq$$
$$\leqq \|S_{n,r}(m,f^*;\ \cdot)\|_{L_p(R)} \leqq (2^r-1)\|f^*\|_{L_p(R)} = (2^r-1)\|f\|_{L_p(D)},$$

which completes the proof.

LEMMA 4.5. *Suppose for* $1\leqq p\leqq\infty$, $f\in L_p(D)$, *for some* $\mu<\min(2r,m)$, $f^{(\mu)}\in$ A.C. $[a,b]$, $[a,b]\subset D$ *and* $f^{(\mu+1)}\in L_p[a,b]$. *Then*

$$\|S_{n,r}(m,f;\ \cdot)-f(\cdot)\|_{L_p[a+Ah_n,b-Ah_n]} \leqq Mh_n^{\mu+1}\|f^{(\mu+1)}\|_{L_p[a,b]} \tag{4.8}$$

where the constants M and A do not depend on f, n, a or b.

In particular, if $f\in L_p(R)$, $1\leqq p\leqq\infty$, $f^{(\mu)}\in$ A.C.$_{\cdot\mathrm{loc}}$ *and* $f^{(\mu+1)}\in L_p(R)$, *then*

$$\|S_{n,r}(m,f;\ \cdot)-f(\cdot)\|_{L_p(R)} \leqq Mh_n^{\mu+1}\|f^{(\mu+1)}\|_{L_p(R)}. \tag{4.9}$$

PROOF. Using the definition of $S_{n,r}(m,f;t)$ and following the proof of Lemma 4.4, we know that if $f(x)=g(x)$ for $x\in(t-Ah_n,t+Ah_n)$, where $A=r(m+1)/2$ for instance, we have $S_{n,r}(m,f;t)=S_{n,r}(m,g;t)$. Using Taylor's formula in L_p and Lemma 4.2, we have

$$\left|S_{n,r}(m,f(\cdot);\ t)-f(t)\right| = \left|S_{n,r}(m,\ R_{\mu+1}(f,\cdot,t);\ t)\right|$$

where

$$R_{\mu+1}(f,u,t) = \frac{1}{\mu!}\int_t^u (u-v)^\mu f^{(\mu+1)}(v)\,dv.$$

Therefore, we have only to estimate

$$\Big\{\int_{a+Ah_n}^{b-Ah_n} \left|S_{n,r}(m,\ R_{\mu+1}(f,\cdot,t);\ t)\right|^p dt\Big\}^{1/p}.$$

Using (3.3), it is enough to estimate

$$\left|\left|S_m^j(m,\ R_{\mu+1}(f,\cdot,t);\ t)\right|\right|_{L_p[a+Ah_n,b-Ah_n]}.$$

Following Lemma 4.4 with $r=1$ (and A there $(m+1)/2$) and choosing A here to be $(r+1)(m+1)/2$, we have

$$\left|\left|S_n^j(m,\ R_{\mu+1}(f,\cdot,t);\ t)\right|\right|_{L_p[a+(r+1)(m+1)h_n/2,\,b-(r+1)(m+1)h_n/2]} \leqq$$
$$\leqq \left|\left|S_n(m,\ R_{\mu+1}(f,\cdot,t);\ t)\right|\right|_{L_p[a+(r-j+2)(m+1)h_n/2,\,b-(r-j+2)(m+1)h_n/2]}.$$

To complete the proof, we have to show that

$$\left|\left|S_n(m,\ R_{\mu+1}(f,\cdot,t);\ t)\right|\right|_{L_p[a+(m+1)h_n,\,b-(m+1)h_n]} \leqq Mh_n^{\mu+1}\|f^{(\mu+1)}\|_{L_p[a,b]}. \tag{4.10}$$

To prove (4.10) we write

$$|S_n(m,\ R_{\mu+1}(f,\ \cdot,\ t),\ t)| =$$

$$\leq \left|\sum \left\{h_n^{-1}\int_0^{h_n}\frac{1}{\mu!}\int_t^{\xi_n+kh_n+u}(v-(\xi_n+kh_n+u))^\mu f^{(\mu+1)}(v)\,dv\,du\right\}\times\right.$$

$$\left.\times G_m\left(-\left(k+\frac{1}{2}\right)+(t-\xi_n)/h_n\right)\right| \leq$$

$$\leq Mh_n^\mu \sum \int_{-mh_n/2}^{mh_n/2}\left|f^{(\mu+1)}\left(v+\xi_n+\left(k+\frac{1}{2}\right)h_n\right)\right|du\, G_m\left(-\left(k+\frac{1}{2}\right)+(t-\xi_n)/h_n\right) \equiv I_n(t).$$

We have to prove now that

$$\|I_n(t)\|_{L_p[a+(m+1)h_n,\, b-(m+1)h_n]} \leq Mh_n^{\mu+1}\|f^{(\mu+1)}\|_{L_p[a,b]}.$$

For $p=\infty$ it is clear as $\sum G_m\left(-\left(k+\frac{1}{2}\right)+(t-\xi)/h\right)=1$. For $p=1$ we use

$$h_n^{-1}\int G_m\left(-\left(k+\frac{1}{2}\right)+(t-\xi_n)/h_n\right)dt = G_m(u)\,du = 1$$

and write

$$\|I_n(t)\|_{L_1[a+(m+1)h_n,\, b-(m+1)h_n]} \leq Mh_n^{\mu+1}\sum{}' \int_{-mh_n/2}^{mh_n/2}\left|f^{(\mu+1)}\left(v+\xi_n+\left(k+\frac{1}{2}\right)h_n\right)\right|dv \leq$$

$$\leq Mh_n^{\mu+1}\|f^{(\mu+1)}\|_{L_1[a,b]}$$

where $\sum'$ is the sum on all k such that

$$(t-\xi)-\left(k+\frac{1}{2}\right)h_n \in [a+(m+1)h_n/2,\ b-(m+1)h_n/2].$$

For other p, that is for $1<p<\infty$, (4.10) follows from combining the above with the Riesz—Thorin Theorem.

Proof of Theorems 3.1 and 3.2. It is well-known that $K_r(f, t^r)_p \sim \omega_r(f, t)_p$ where

$$K_r(f, t^r)_p = \inf_{g^{(r-1)}\in \text{A.C.}} \{\|f-g\|_{L_p[a,b]} + t^r\|g^{(r)}\|_{L_p[a,b]}. \tag{4.11}$$

The results follow now from Lemma 4.4 and 4.5, writing $f=f-g+g$ where $\|f-g\|_p \leq M\omega_r(f, h_n)_p$ and $\|g^{(r)}\|_p \leq Mh_n^{-r}\omega_r(f, h_n)_p$ in the appropriate intervals.

5. The main result in $L_p(R)$ under an a priori condition

We are now able to prove the main result of the paper under the a priori assumption that f belongs to some Lipschitz class. It will be in the next section that this a priori condition is dropped for $p>1$.

THEOREM 5.1. *Suppose* $f\in L_p(R)$, $1\leq p\leq\infty$, $\|\Delta_h f\|_p\leq Mh^\delta$ *for some* $\delta>0$ *and* $\varphi\in\Phi$. *Then*

(5.1) $$\|\Delta^r a_k(n)\|_{l_p}=O\big(\varphi(h_n)h_n^{-1/p}\big)\Leftrightarrow\omega_r(f,t)_p=O\big(\varphi(t)\big)$$

and

(5.2) $$\|\Delta^r a_k(n)\|_{l_p}\sim\varphi(h_n)h_n^{-1/p}\Leftrightarrow\omega_r(f,t)_p\sim\varphi(t).$$

Recall $\varphi\in\Phi$ if $\varphi(h)$ is an increasing function for $h\geq 0$, satisfying $\lim\limits_{h\to 0+}\varphi(h)=0$ and for some η there exists $L=L(\eta)$ such that $\varphi(h)/\varphi(Lh)<\eta<1$ for all $h>0$.

PROOF. We first prove (5.1). To demonstrate the implication "$\Leftarrow$" we write

$$\|\Delta^r a_k(n)\|_{l_p}=\Big\{\sum_k\Big|\frac{1}{h_n}\int_0^{h_n}\vec{\Delta}^r_{h_n}f(\xi_n+kh_n+u)\,du\Big|^p\Big\}^{1/p}\leq$$

$$\leq\Big\{\sum_k\frac{1}{h_n}\int_0^{h_n}|\vec{\Delta}^r_{h_n}f(\xi_n+kh_n+u)|^p\,du\Big\}^{1/p}=h_n^{-1/p}\Big\{\int_R|\vec{\Delta}^r_{h_n}f(u)|^p\,du\Big\}^{1/p}\leq h_n^{-1/p}\omega_r(f,h_n)_p.$$

We have to prove (5.1) in the other direction now. For this we use the approximation operator defined in (3.3) (see also (3.2) and (3.4)) and write

$$\begin{aligned}\|\Delta^r_h f(x)\|_{L_p}&\leq\|\Delta^r_h(f-S_{n,r}(f))\|_p+\|\Delta^r_h S_{n,r}(f)\|_p\leq\\&\leq 2^r M\omega_r(f,h_n)_p+(2^r-1)\sup_{1\leq j\leq r}\|\Delta^r_h S^j_n(f)\|_p.\end{aligned}$$

To estimate $\|\Delta^r_h S^j_n(f)\|_p$ we write, using (4.3),

$$\|\Delta^r_h S^j_n(f)\|^p=Ch^r\Big\|\Big(\frac{d}{dx}\Big)^r S^j_n(f)\Big\|_p=Ch^r\big\|h_n^{-r}S_n(m-r,\vec{\Delta}^r_{h_n}S^{j-1}_n(f);\ \cdot)\big\|_p.$$

Using the fact that S_n is a contraction (which is shown in Lemma 4.3, $r=1$), we have

(5.3) $$\|\Delta^r_h S^j_n(f)\|_p\leq C\Big(\frac{h}{h_n}\Big)^r\|\vec{\Delta}^r_{h_n}S^{j-1}_n(f)\|_p\leq\ldots\leq C^{j-1}h^r\Big\|\Big(\frac{d}{dx}\Big)^r S_n(f)\Big\|_p.$$

Utilizing again the method of Lemma 4.3, we obtain

$$\Big\|\Big(\frac{d}{dx}\Big)^r S_n(f)\Big\|_\infty\leq h_n^{-r}\|\Delta^r a_k(n)\|_{l_\infty}$$

and

$$\Big\|\Big(\frac{d}{dx}\Big)^r S_n(f)\Big\|_1\leq h_n^{-r}\|\Delta^r a_k(n)\|_{l_1}h_n$$

which together with the Riesz—Thorin interpolation theorem yield

$$\left\|\left(\frac{d}{dx}\right)^r S_n(f)\right\|_p \leqq h_n^{-r}\|\Delta^r a_k(n)\|_{l_p} h_n^{1/p}.$$

Therefore, we have

$$\|\Delta_h^r f(x)\|_p \leqq 2^r M\omega_r(f, h_n)_p + C(1)(h/h_n)^r \|\Delta^r a_k(n)\|_{l_p} h_n^{1/p}. \tag{5.4}$$

Without loss of generality we may assume the sequence h_n satisfies $1 < T_1 \leqq h_n/h_{n+1} \leqq T_2 < \infty$ where T_1 and T_2 are large enough. (We will choose T_i later.) We have (following (5.4)) for $h_n \leqq h < h_{n-1}$

$$\|\Delta_h^r f\|_p \leqq 2^r M\omega_r(f, h_n)_p + C(1)(h/h_n)^r \varphi(h_n) \leqq M^{s+1}(2^r)^{s+1}\omega_r(f, h_{n+s})_p +$$

$$+ C(1)(h/h_n)^r\varphi(h_n) + C(1)\sum_{k=1}^{s}(h_{n+k-1}/h_{n+k})^r(2^r)^k M^k\varphi(h_{n+k}) \leqq$$

$$\leqq (2^r)^{s+1}M^{s+1}\omega_r(f, h_{n+s})_p + C(1)(h/h_n)^r\varphi(h_n) + \sum_{k=1}^{s} C(1)T_2^r(2^r)^k M^k\varphi(h_{n+k}).$$

We choose $T_1 \geqq L^m$ such that $M2^r\eta^m \leqq q < 1$ and

$$\varphi(h) \leqq \eta^m\varphi(T_1 h),$$

and therefore,

$$\varphi(h_{n+k}) \leqq \eta^{mk}\varphi(h_n).$$

This implies

$$\|\Delta_h^r f\|_p \leqq M^{s+1}(2^r)^{s+1}\omega_r(f, h_{n+s})_p + C(2)\varphi(h_n).$$

For $f \in \mathrm{Lip}\,\delta$ we can choose T_1 so that $M^{s+1}(2^r)^{s+1}\omega_r(f, h_{n+s}) = o(1)$, $s \to \infty$ and this completes the proof of (5.1).

To prove (5.2) we have only to show that $\omega_r(f, t)_p \sim \varphi(t)$ implies $\|\Delta^r a_k(n)\|_{l_p} \geqq M_1\varphi(h_n)h_n^{-1/p}$ (as the rest was proved in (5.1) or is obvious).

We use now:

(a) $\omega_r(f, t)_p \sim \varphi(t)$, that is, $K^{-1}\varphi(t) \leqq \omega_r(f, t)_p \leqq K\varphi(t)$;
(b) $\|f - S_{n,r}(f)\|_p \leqq M\omega_r(f, h_n)_p$ where M is independent of f and h_n; and
(c) $\varphi(h) \leqq \eta\varphi(Lh)$ where L depends on η but not on h.

We choose a constant B such that $B = L^m$ and $2^r MK^2\eta^m < 1/2$, and for $Bh_n < t$ we have

$$2^r M\omega_r(f, h_n)_p \leqq 2^r MK\varphi(h_n) \leqq 2^r MK\eta^m\varphi(t) \leqq 2^r MK^2\eta^m\omega_r(f, t)_p \leqq \frac{1}{2}\omega_r(f, t)_p.$$

Using (5.4), we obtain for $h_n < t/B \leqq h_{n-1}$

$$\omega_r(f, t)_p \leqq 2^r M\omega_r(f, h_n)_p + C(1)\sup_{h \leqq t}(h/h_n)^r\|\Delta^r a_k(n)\|_p h_n^{1/p}, \tag{5.5}$$

and therefore,

$$\omega_r(f, t)_p \leqq C(2)\|\Delta^r a_k(n)\|_p h_n^{1/p} \quad \text{for} \quad h_n \leqq t/B \leqq h_{n-1}.$$

This implies

$$\omega_r(f, h_n)_p \leqq C(3)\|\Delta^r a_k(n)\|_p h_n^{1/p},$$

which completes the proof of our theorem.

6. Relaxing the a priori condition

In Section 5 the condition $f\in \mathrm{Lip}\,\delta$, that is $\|\Delta_h^r f\|_p \leqq Mh^\delta$ for some $\delta>0$, was imposed. In fact it was shown in [8] that for $p=\infty$ and $\varphi(t)=t^\alpha$ this a priori condition is unnecessary. We will show here that it is superfluous for $1<p\leqq\infty$.

In fact, all we need to show is the following lemma.

LEMMA 6.1. *For* $1<p\leqq\infty$, $f\in L_p(R)$ *and* $a_k(n)$ *defined in* (2.1)

$$\|\Delta^r a_k(n)\|_{l_p} h_n^{1/p} \leqq M h_n^\beta$$

implies

$$\omega_1(f, h)_p \equiv \omega(f, h)_p \leqq M_1 h^{\beta_1} \tag{6.1}$$

for some β_1 *and* M_1 *dependent on* β *and* M *but not on* f. (We may choose $\beta_1=\dfrac{\beta}{2r}\dfrac{p-1}{p}\Big)$.

That Lemma 6.1 is sufficient follows from the fact that $\varphi(t)$ treated in the last section satisfied $\varphi(h)\leqq\eta\varphi(Lh)$ for some $L>1$, and this implies for $h_0/L^n < t\leqq h_0/L^{n-1}$ that

$$\varphi(t) \leqq \eta^{n-1}\varphi(h_0) \leqq (L^{-\beta})^{n-1}\varphi(h_0) \leqq \big(\varphi(h_0)/h_0^\beta\big)(h_0/L^{n-1})^\beta \leqq \big(\varphi(h_0)/h_0^\beta\big)L^\beta t^\beta \leqq Ct^\beta$$

for β satisfying $L^{-\beta}<\eta$.

We will show in Lemma 6.2 that for $p=1$ (and in fact for other p as well) (b) of Theorem 2.2 can replace the a priori condition.

PROOF of Lemma 6.1. We first recall [7, p. 308] that for $1\leqq p\leqq\infty$ (including $p=1$) the inequalities

$$\|\Delta^r a_k(n)\|_{l_p} \leqq M(r)h_n^{\beta-1/p} \quad\text{and}\quad \|a_k(n)\|_{l_p} \leqq M(0)h_n^{-1/p}$$

imply

$$\|\Delta a_k(n)\|_{l_p} \leqq M(1)h_n^{\frac{\beta}{r}-\frac{1}{p}}. \tag{6.2}$$

It will suffice to show now that (6.2) implies (6.1). We will use (6.2) on a subsequence h_{n_i} satisfying $1<T_1\leqq h_{n_i}/h_{n_{i+1}}\leqq T_2<\infty$ and with no loss of generality we assume that h_n is already that subsequence, that is $1<T_1\leqq h_n/h_{n+1}\leqq T_2<\infty$. We will define f_n for which

$$\|f-f_n\|_p \leqq Kh_n^{\beta_1} \tag{6.3}$$

and for $h_n\leqq h<h_{n-1}$

$$\|f_n(x)-f_n(x+h)\|_p \leqq Kh_n^{\beta_1}. \tag{6.4}$$

The inequalities (6.3) and (6.4) will imply

$$\|f(x)-f(x+h)\|_p \leqq \|f(x)-f_n(x)\|_p+\|f(x+h)-f_n(x+h)\|_p+$$
$$+\|f_n(x)-f_n(x+h)\|_p \leqq 3Kh_n^{\beta_1}.$$

Given $\{h_k\}_{k=n}^{\infty}$, we can select a subsequence $\{\eta_k\}_{k=n}^{\infty}$ such that $\eta_n=h_n$ and $\eta_l=h_{n_l}$ (for each n a new subsequence η_k) and write

$$(6.5)\qquad f_{\eta_l}(t)=\frac{1}{\eta_l}\int_0^{\eta_l} f(\xi_{n_l}+k\eta_l+u)\,du \quad \text{for} \quad t-\xi_{n_l}-k\eta_l\in(0,\eta_l)$$

which can be written also as

$$(6.6)\qquad f_{\eta_l}(t)=S_{n_l}(1,f;\,t)\equiv\sum a_k(n_l)G_1\left(-\left(k+\frac{1}{2}\right)+(t-\xi_{n_l})/h_{n_l}\right).$$

Obviously, $f_n(t)=f_{\eta_n}(t)$. We will choose the subsequence η_l in such a way that

$$(6.7)\qquad \|f_{\eta_l}-f_{\eta_{l+1}}\|_p \leqq K\eta_l^{\beta_1}.$$

Recall that as η_n is a subsequence of h_n which satisfies $1<T_1<h_n/h_{n+1}$, we have also $1<T_1<\eta_l/\eta_{l+1}$. To show that (6.3) follows from (6.7) we need to show that for $1\leqq p<\infty$

$$(6.8)\qquad \lim_{n\to\infty}\|f-f_n\|_p=0 \quad \left(\text{in particular } \lim_{l\to\infty}\|f-f_{\eta_l}\|_p=0\right),$$

and for $p=\infty$ (actually for $1\leqq p\leqq\infty$)

$$(6.8)'\qquad \lim_{n\to\infty}f_n(x)=f(x) \text{ a.e.} \quad \left(\text{in particular } \lim_{l\to\infty}f_{\eta_l}(x)=f(x) \text{ a.e.}\right).$$

To prove (6.8), we write for $1\leqq p<\infty$

$$\|f(t)-f_n(t)\|_p^p=\sum_k\int_0^{h_n}\left|f(\xi_n-kh_n+v)-h_n^{-1}\int_0^{h_n}f(\xi_n+kh_n+u)\,du\right|^p dv \leqq$$
$$\leqq \sum_k h_n^{-1}\int_0^{h_n}\int_0^{h_n}|f(\xi_n+kh_n+v)-f(\xi_n+kh_n+u)|^p\,du\,dv \leqq$$
$$\leqq \sum_k h_n^{-1}\int_{-h_n}^{h_n}\int_0^{h_n}|f(\xi_n+kh_n+v)-f(\xi_n+kh_n+v+w)|^p\,dv\,dw \leqq$$
$$\leqq 2\omega(f,h_n)_p^p=o(1) \qquad n\to\infty.$$

For $p=\infty$ (6.8)′ is easy to show.

We will now prove (6.4) for $h_n\leqq h<h_{n-1}$. For $0\leqq\tau<h_n$ we have $[(\tau+h)/h_n]=m$ for $0\leqq\tau<\alpha h_n$ and $[(\tau+h)/h_n]=m+1$ for $\alpha h_n\leqq\tau<h_n$ where $0\leqq\alpha=\alpha(h)<1$ is fixed and $m<M+1$ where $h_n/h_{n+1}\leqq M$. We can now write

$$\|f_n(t)-f_n(t+h)\|_p^p = \sum_k \Big\{\int_0^{\alpha h_n} \Big|h_n^{-1}\int_0^{h_n} \big(f(\xi_n+kh_n+v)-f(\xi_n+(k+m)h_n+v)\big)\,dv\Big|^p +$$

$$+\int_{\alpha h_n}^{h_n} \Big|h_n^{-1}\int_0^{h_n} f(\xi_n+kh_n+v)-f\big(\xi_n+(k+m+1)h_n+v\big)\,dv\Big|^p\Big\} \leqq$$

$$\leqq \alpha h_n \sum_k |a_n(k+m)-a_n(k)|^p+(1-\alpha)h_n\sum_k |a_n(k+m+1)-a_n(k)|^p \leqq$$

$$\leqq \max_{j=m,m+1} h_n \sum |a_n(k+j)-a_n(k)|^p \leqq (m+1)^p\|\Delta a_n(k)\|_{l_p}^p h_n,$$

which implies (6.4) with $\beta_1=\dfrac{\beta}{r}$ or $0<\beta_1\leqq\dfrac{\beta}{r}$.

The main step of the proof, and the only remaining step, is establishing (6.7), that is, finding a subsequence of $\{h_k\}_{k=n}^{\infty}$, $\{\eta_\kappa\}_{k=n}^{\infty}$ for which $h_n=\eta_n$, and $f_{\eta_\kappa}(x)$ satisfies

$$\|f_{\eta_l}(x)-f_{\eta_{l+1}}(x)\|_p \leqq K_1\eta_l^{\beta_1}$$

for some $\beta_1>0$. $\left(\beta_1=\dfrac{\beta}{2r}\dfrac{p-1}{p}\right.$ for $p<\infty$ will do.$\Big)$ We now select the subsequence $\{\eta_k\}_{k=n}^{\infty}$ of $\{h_k\}_{k=n+1}^{\infty}$. Recall that we already have $1<T_1<h_l/h_{l+1}\leqq T_2<\infty$. We choose η_{l+1} (once η_l is given) so that

$$T_2^{-1}\eta_l^{1+\beta_2}<\eta_{l+1}\leqq\eta_l^{1+\beta_2} \quad\text{where}\quad \beta_2=\frac{\beta}{2r} \quad\text{for instance.}$$

We define $f^*_{\eta_{l+1}}(x)$ by

$$\text{(6.9)}\qquad f^*_{\eta_{l+1}}(x)=\frac{1}{s\eta_{l+1}}\int_0^{s\eta_{l+1}} f(\xi_{n_{l+1}}+k\eta_{l+1}+u)\,du$$

for x satisfying $x-\xi_{\eta_{l+1}}-k\eta_{l+1}\in[0,s\eta_{l+1}]$ and where $k=k(m)$ and $s=s(m)$ are given for any m by

$$k(m)=\max\,(j\colon \xi_{n_{l+1}}+j\eta_{l+1}\leqq\xi_{n_l}+m\eta_l\}$$

and

$$s=s(m)=\max\big(j-k(m)\colon \xi_{n_{l+1}}+j\eta_{l+1}\leqq\xi_{n_l}+(m+1)\eta_l\big\}.$$

It is clear from the choice of η_l that $(\xi_{\eta_{l+1}}+k\eta_{l+1},\xi_{\eta_{l+1}}+(k+s)\eta_{l+1})$ are disjoint and that $s=s(m)\leqq M\eta_l^{-\beta_2}$. In fact, for a given l, $s(m)$ can assume only two consecutive values.

The estimate of $f_{\eta_l}-f_{\eta_{l+1}}$ will be divided into two parts, the estimate of $f_{\eta_l}-f^*_{\eta_{l+1}}$ and that of $f^*_{\eta_{l+1}}-f_{\eta_{l+1}}$. We first estimate $f_{\eta_{l+1}}-f^*_{\eta_{l+1}}$ writing

$$\|f_{\eta_{l+1}}-f^*_{\eta_{l+1}}\|_p^p\leqq\sum_m\sum_{j=k(m)}^{k(m)+s(m)-1}\|f_{\eta_{l+1}}-f^*_{\eta_{l+1}}\|^p_{L_p[\xi_{n_{l+1}}+j\eta_{l+1},\,\xi_{n_{l+1}}+(j+1)\eta_{l+1}]}.$$

We now estimate for $k(m) \leqq j < k(m)+s(m)$

$$\|f_{\eta_{l+1}}-f^*_{\eta_{l+1}}\|^p_{L_p[\xi_{n_{l+1}}+j\eta_{l+1},\,\xi_{n_{l+1}}+(j+1)\eta_{l+1}]} =$$

$$= \eta^p_{l+1}\left|\frac{1}{s}\sum_{i=k(m)}^{k(m)+s(m)-1} a_i(n_{l+1})-a_j(n_{l+1})\right|^p =$$

$$= \eta^p_{l+1}\frac{1}{s^p}\left|\sum_{i=k(m)}^{k(m)+s(m)-1} |a_i(n_{l+1})-a_j(n_{l+1})|\right|^p \leqq$$

$$\leqq \eta^p_{l+1}\left|\sum_{i=k(m)}^{k(m)+s(m)-1} |\Delta a_i(n_{l+1})|\right|^p \leqq \eta^p_{l+1}s(m)^p\frac{1}{s(m)}\sum_{i=k(m)}^{k(m)+s(m)-1} |\Delta a_i(n_{l+1})|^p.$$

Therefore,

$$\|f_{\eta_{l+1}}-f^*_{\eta_{l+1}}\|_p \leqq \eta_{l+1}\Big\{\sum_m s(m)^p\sum_{i=k(m)}^{k(m)+s(m)-1} |\Delta a_i(n_{l+1})|^p\Big\}^{1/p} \leqq$$

$$\leqq M\eta_{l+1}\eta_l^{-\beta_2}\{\sum |a_i(n_{l+1})|^p\} \leqq M_1\eta_l\|\Delta a_{n_{l+1}}\|_p \leqq M_2\eta_l(\eta_{l+1})^{\frac{\beta}{r}-\frac{1}{p}} \leqq$$

$$\leqq M_3\eta_l\eta_l^{(1+\beta_2)\left(\frac{\beta}{r}-\frac{1}{p}\right)}.$$

The last expression is a positive power of η_l and if β_2 is chosen appropriately, for instance $\beta_2 \leqq \frac{\beta}{2r}$, it will imply $\|f_{\eta_{l+1}}-f^*_{\eta_{l+1}}\| \leqq M\eta_l^{\beta/2r}$. That is, $\beta_1=\frac{\beta}{2r}$ is appropriate for this part of the proof.

We now estimate $f_{\eta_l}-f^*_{\eta_{l+1}}$, and in fact, this is the only place in the proof where the condition $p>1$ is used. (We will not treat $p=\infty$ as it follows similar and easier lines and also was proved in [8], though the proof there is not easier.)

$$\|f_{\eta_l}-f^*_{\eta_{l+1}}\|^p_p \leqq \sum_m \big(\xi_{\eta_{l+1}}+\big(k(m)+s(m)\big)\eta_{l+1}-\xi_{n_l}-m\eta_l\big)\times$$

$$\times\left|\frac{1}{s\eta_{l+1}}\int_0^{s(m)\eta_{l+1}} f(\xi_{n_{l+1}}+k(m)\eta_{l+1}+u)\,du-\frac{1}{\eta_l}\int_0^{\eta_l} f(\xi_{n_l}+m\eta_l+u)\,du\right|^p +$$

$$+\sum_m (\xi_{n_l}+m\eta_l-\xi_{n_{l+1}}-k(m)\eta_{l+1})\left|\frac{1}{s\eta_{l+1}}\int_0^{s(m)\eta_{l+1}} f(\xi_{n_{l+1}}+k(m)\eta_{l+1}+u)\,du-\right.$$

$$\left.-\frac{1}{\eta_l}\int_0^{\eta_l} f(\xi_{n_l}+(m-1)\eta_l+u)\,du\right|^p \leqq$$

$$\leqq \sum_m s(m)\eta_{l+1}I(l,m)+\sum_m \eta_{l+1}2^p(I(l,m)+|a_m(n_l)-a_{m-1}(\eta_l)|^p)$$

where $I(l, n)$ is given by

$$I(l,n)=\left|\frac{1}{s\eta_{l+1}}\int_0^{s\eta_{l+1}} f(\xi_{n_{l+1}}+k(m)\eta_{l+1}+u)\,du-\frac{1}{\eta_l}\int_0^{\eta_l} f(\xi_{n_l}-m\eta_l+u)\,du\right|^p.$$

To estimate $I(l, m)$ we write

$$I(l, m) \leqq \left(\left|\left(\frac{1}{s\eta_{l+1}} - \frac{1}{\eta_l}\right) \int\limits_{\xi_{n_l}+m\eta_l-\xi_{n_{l+1}}-k(m)\eta_{l+1}}^{s(m)\eta_{l+1}} f(\xi_{n_{l+1}} + k(m)\eta_{l+1} + u)\, du\right| + \right.$$

$$+ \left|\frac{1}{s\eta_{l+1}} \int\limits_0^{\xi_{n_l}+m\eta_l-\xi_{n_{l+1}}-k(m)\eta_{l+1}} f(\xi_{n_{l+1}} + k(m)\eta_{l+1} + u)\, du\right| +$$

$$\left. + \left|\frac{1}{\eta_l} \int\limits_{\xi_{n_{l+1}}+(k(m)+s(m))\eta_{l+1}-\xi_{n_l}-m\eta_l}^{\eta_l} f(\xi_{n_l} + m\eta_l + u)\, du\right|\right)^p \equiv$$

$$\equiv (J_1(m) + J_2(m) + J_3(m))^p \leqq 3^p(J_1^p + J_2^p + J_3^p).$$

To estimate J_2^p we write

$$J_2^p \leqq \frac{1}{(s\eta_{l+1})^p} (\xi_{n_l} + m\eta_l - \xi_{n_{l+1}} - k(m)\eta_{l+1})^{p-1} \int\limits_0^{s(m)\eta_{l+1}} |f(\xi_{n_{l+1}} + m\eta_{l+1} + u)|^p\, du \leqq$$

$$\leqq \frac{1}{s^p \eta_{l+1}} \int\limits_0^{s(m)\eta_{l+1}} |f(\xi_{n_{l+1}} + k\eta_{l+1} + u)|^p\, du.$$

Similarly,

$$J_3^p \leqq \frac{1}{\eta_l^p} (\eta_{l+1})^{p-1} \int\limits_0^{\eta_l} |f(\xi_{\eta_l} + \eta_l + u)|^p\, du.$$

To estimate J_1^p we write

$$J_1^p \leqq \left(\frac{\eta_l - s\eta_{l+1}}{s\eta_{l+1}\eta_l}\right)^p (s\eta_{l+1} - \xi_{n_l} - m\eta_l + \xi_{n_{l+1}} + k(m)\eta_{l+1})^{p-1} \times$$

$$\times \int\limits_0^{s\eta_{l+1}} |f(\xi_{n_{l+1}} + k(m)\eta_{l+1} + u)|^p\, du \leqq$$

$$\leqq \left(\frac{\eta_{l+1}}{\eta_l}\right)^p \frac{1}{s\eta_{l+1}} \int\limits_0^{s\eta_{l+1}} |f(\xi_{n_{l+1}} + k(m)\eta_{l+1} + u)|^p\, du.$$

We now write

$$\|f_{\eta_{l+1}} - f_{\eta_l}^*\|_p^p \leqq \sum_m 3^p(s + 2^p)\eta_{l+1}(J_1(m)^p + J_2(m)^p + J_3^p(m)) + \sum_m{}' 2^p\eta_{l+1}|\Delta a_{m-1}(n_l)|^p \leqq$$

$$\leqq M \sum_m \eta_l^{-\beta_2}\eta_{l+1} \left(\frac{\eta_{l+1}}{\eta_l}\right)^p \frac{1}{s\eta_{l+1}} \int\limits_0^{s\eta_{l+1}} |f(\xi_{n_{l+1}} + k\eta_{l+1} + u)|^p\, du +$$

$$+ M \sum_m \frac{s\eta_{l+1}}{s^p\eta_{l+1}} \int\limits_0^{s(m)\eta_{l+1}} |f(\xi_{n_{l+1}} + k\eta_{l+1} + u)|^p\, du +$$

$$+ M \sum_m \frac{\eta_{l+1}^{p-1}}{\eta_l^p} s\eta_{l+1} \int\limits_0^{\eta_l} |f(\xi_{n_l} + m\eta_l + u)|^p\, du + 2^p\eta_{l+1}\|\Delta a_m(n_l)\|_{l_p}^p \leqq$$

$$\leqq M_1[\eta_l^{\beta_2 p}\|f\|_p^p + \eta_l^{\beta_2(p-1)}\|f\|_p^p + \eta_l^{\beta_2(p-1)}\|f\|_p^p + \eta_l^{-1+\beta p/r}\eta_l^{1+\beta_2}].$$

Therefore, it is easy to see for $\beta>0$ and $p>1$ that

$$\|f_{\eta_{l+1}}-f^*_{\eta_{l+1}}\| \leqq M[\eta_l^{\beta_2(p-1)/p}\|f\|+\eta_l^{(\beta/r)+(\beta_2/p)}],$$

which completes the proof. That is, for $\beta_2=\frac{\beta}{2r}$, $\beta_1=\frac{\beta}{2r}\frac{p-1}{p}$ $\left(\text{for } p=\infty,\ \beta_1=\frac{\beta}{2r}\right)$.

LEMMA 6.2. *Suppose $\xi_n=\xi$ and $h_n=m(n)h_{n+1}$ where $m(n)$ is an integer satisfying $1<m(n)\leqq M$, and let $\varphi\in\Phi$ be as given in Theorem 2.1. Then*

(6.10) $$\|\Delta^r a_k(n)\|_{l_p(R)} = O(\varphi(h_n)h_n^{-1/p}) \Rightarrow \omega_r(f,t)_p = O(t^{\beta_1}).$$

PROOF. Using an earlier argument, we have $\|\Delta^r a_k(n)\|_{l_p(R)}=O(h_n^{\beta-1/p})$, and therefore, $\|\Delta a_k(n)\|_{l_p(R)}=O(h_n^{\frac{\beta}{r}-\frac{1}{p}})$. The proof will be complete now if we define f_n as in the proof of Lemma 6.1 by (6.5) where $h_n=\eta_n$. The estimate of $\|f_n-f_{n+1}\|$ which replaces (6.7) here is reduced to $\|f^*_{n+1}-f_{n+1}\|$ as

$$f_n(x)=f^*_{n+1}=\frac{1}{m(n)h_{n+1}}\int_0^{m(n)h_n} f(\xi+kh_{n+1}+u)\,du$$

for x satisfying $x-\xi-kh_{n+1}\in[0,m(n)h_{n+1}]$ where k is divisible by $m(n)$. The estimate of $\|f^*_{n+1}-f_{n-1}\|_p$ follows closely the estimate of $\|f^*_{\eta_{l+1}}-f_{\eta_{l+1}}\|_p$ in the proof of Lemma 6.1 which did not depend on $p>1$. Consequently, we obtain $\|f_n-f_{n+1}\|\leqq$ $\leqq M\eta^{\beta_1}$ where $\beta_1=\frac{\beta}{2r}$ if $\|\Delta a_k(n)\|_{l_p(R)}=O(h_n^{(\beta/r)-(1/p)})$.

7. Intervals with a finite end point

To extend our result to include $[a,b]$ or $[a,\infty)$ where a and b are finite we will use the concept of main-part moduli of smoothness (see [10, Ch. 3]). That is, the method will differ from the method in [8] where a Whitney-type extension result for discrete data was proved and used. The present proof is easier and yields more general results but uses a deep result from [10]. The method will imply also an extension of the result of [8] which together with other extensions will be described in Section 8.

With no loss of generality we may assume that the domain is $[0,1]$ or $[0,\infty)$. The main-part modulus of smoothness is given for $[0,1]$ by

(7.1) $$\Omega_r(f,t)_p \equiv \sup_{0<h\leqq t}\|\Delta_h^r f\|_{L_p[rh,1-rh]}$$

and for $[0,\infty)$ by

(7.2) $$\Omega_r(f,t)_p \equiv \sup_{0<h\leqq t}\|\Delta_h^r f\|_{L_p[rh,\infty]}.$$

We recall that $\Delta_h^r f=\sum_{k=0}^{r}(-1)^k\binom{r}{k}f\left(x+\frac{rh}{2}-kh\right)$ for $\left[x-\frac{rh}{2},x+\frac{rh}{2}\right]\subset D$ where $D=[0,1]$ (in case of (7.1)) or $D=[0,\infty)$ (in case of (7.2)) and $\Delta_h^r f=0$ otherwise.

The result which we will use was proved in [10] and is given in the present special case by

LEMMA 7.1. *For* $\Omega_r(f,t)_p$ *and* $\omega_r(f,t)$ *given in* (7.1), (7.2) *and* (1.1) *we have*

$$\Omega_r(f,t)_p \leqq \omega_r(f,t)_p \equiv \int_0^t \big(\Omega_r(f,\tau)/\tau\big)\,d\tau. \tag{7.3}$$

The result in [10] was proved in a wider context for a step weight function $\varphi(x)$ satisfying some conditions; the definitions (7.1) and (7.2) and the result (7.3) apply [10] to the step-weight function $\varphi(x)=1$. Examining for $p=\infty$ the function $f(x)=(\log x)^{-1}$ for $0<x\leqq 1/2$ and $f(x)\in C^\infty$ otherwise, we have $\Omega_1(f,t)_\infty=O((\log t)^{-2})$ and $\omega_1(f,t)_\infty=O(|\log t|^{-1})$. For $\Omega_r(f,t)_p=O(t^\alpha)$ we have $\omega_r(f,t)_p=O(t^\alpha)$. We now recall for $a_k(n)$ given in (2.1)

$$\|\Delta^r a_k(n)\|_{l_p(D)} = \{\textstyle\sum' |\Delta^r a_k(n)|^p\}^{1/p}$$

where $\sum'_k$ signifies that for a given n the sum is on all k satisfying

$$[\xi_n+kh_n,\ \xi_n+(k+r+1)h_n]\subset D.$$

We are now able to prove Theorems 2.1 and 2.2 for the domains $[0, 1]$ and $[0, \infty)$.

PROOF of Theorems 2.1 and 2.2 for $D=[0,1]$ and $D=[0,\infty)$. The proof follows the line of argument of the proof of the theorems for $D=R$ with some modifications.

We first show that for $\varphi\in\Phi$, as restricted in Theorem 2.1 and $1<p\leqq\infty$ or $p=1$ with ξ_n and h_n satisfying condition (b) of Theorem 2.2, we have

$$\|\Delta^r a_k(n)\|_{l_p(D)} = O\big(\varphi(h_n)h_n^{-1/p}\big) \Rightarrow \|\Delta_h f\|_{L_p(D(h))} = O(h^\delta) \tag{7.4}$$

for some $\delta>0$ where $D(h)=[h,\infty)$ if $D=[0,\infty)$, and $D(h)=[h, 1-h]$ if $D=[0,1]$. As $\Omega(f,t)_p=\sup_{0\leqq h\leqq t}\|\Delta_h f\|_{L_p(D(h))}=O(t^\delta)$, (7.3) implies $\omega(f,t)_p=O(t^\delta)$.

To prove (7.4) we demonstrate that $\|\Delta^r a_k(n)\|_{l_p(D)}\leqq Mh^{\beta_2-1/p}$ and $\|a_k(n)\|_{l_p(D)}\leqq Mh_n^{-1/p}$ implies

$$\|\Delta a_k(n)\|_{l_p(D)} \leqq Mh_n^{(\beta_3/r)-(1/p)} \quad\text{where}\quad \beta_3=\min(\beta_2, 1).$$

In fact, for $D=R_+$ we may write $\beta_3=\beta_2$ and the above choice of β_3 is necessary for the proof only in case $D=[0,1]$. For $D=R_+$ we use the result for $l_p(N)$ in [7] (rather than the result for $l_p(Z)$ which we used earlier) and obtain the implication.

To prove (7.4) for $D=[0,1]$, we define the sequence $b_k(n)$ by

$$\begin{aligned} &b_k(n)=1 \quad\text{for}\quad \xi_n+kh_n\leqq\frac{1}{3}\quad\text{or}\quad k\leqq k_1\\ &b_k(n)=0 \quad\text{for}\quad \xi_n+kh_n\geqq\frac{1}{3}\quad\text{or}\quad k\geqq k_2\\ &b_k(n)=1-(k-k_1)/(k_2-k_1)\quad\text{for}\quad k_1<k<k_2. \end{aligned} \tag{7.5}$$

Considering $a_k(n)=a_k(n)b_k(n)+a_k(n)(1-b_k(n))$, it is enough to estimate $\|\Delta(a_k(n)b_k(n))\|_{l_p(D)}$ and $\|\Delta(a_k(n)(1-b_k(n)))\|_{l_p(D)}$. As the estimates of

$$\|\Delta(a_k(n)b_k(n))\|_{l_p(D)} \quad \text{and} \quad \|\Delta(a_k(n)(1-b_k(n)))\|_{l_p(D)}$$

are symmetric, we estimate only the first. We observe that

$$\|\Delta^r(a_k(n)b_k(n))\|_{L_p(D)} \leqq$$

$$\leqq \|\Delta^r a_k(n)\|_{l_p(D)}\|b_k(n)\|_{l_\infty(D)}+C\|a_k(n)\|_{l_p(D)}\|\Delta b_k(n)\|_{l_\infty(D)} \leqq$$

$$\leqq Mh^{\beta_2-1/p}+C_1h_n^{-1/p}h_n \leqq C_2h_n^{\min(\beta_2,1)-1/p}$$

as (7.5) implies $\|b_k(n)\|_{l_\infty(D)}=1$ and $\|\Delta b_k(n)\|_{l_\infty(D)}=O(h_n)$. We now look at the sequence $c_k(n)=a_k(n)b_k(n)$ if $\xi+kh_n<1$ and $c_k(n)=0$ otherwise, and use the result on $l_p(N)$ to obtain the estimate on $\|\Delta c_k(n)\|_{l_p}=\|\Delta(a_k(n)b_k(n))\|_{l_p(D)}$.

The arguments of Section 6 now imply $\|\Delta_h f\|_{L_p(D(h))}=O(h^\delta)$ which can be written as $\Omega(f,t)_p=O(t^\delta)$, and that implies, via (7.3), $\omega(f,t)_p=O(t^\delta)$.

The main part of the proof is the verification of the implication "$\Rightarrow$" of (2.5). (The implication "$\Leftarrow$" of (2.5) follows exactly the proof of the corresponding implication of (5.1) in Theorem 5.1.) We first show that as $1<T_1\leqq h_n/h_{n+1}<T_2$ where T_1 is as chosen in the proof of Theorem 5.1 and $h_n\leqq Bh$, where B is to be chosen later, we have for $D=[0,1]$

$$\|\Delta_h^r f\|_{L_p[hr,1-hr]} \leqq C\sup_{0<h\leqq h_n}\|\Delta_h^r f\|_{L_p[rh_n,1-rh_n]}+C\left(\frac{h}{h_n}\right)^r\varphi(h_n) \tag{7.6}$$

and a similar result (where $[hr,\infty)$ replaces $[hr,1-hr]$) for $D=R_+$. To prove (7.6) we write

$$\|\Delta_h^r f\|_{L_p[hr,1-hr]} \leqq \|\Delta_h^r(f-S_{n,r}(f))\|_{L_p[hr,1-hr]}+\|\Delta_h^r S_{n,r}(f)\|_{L_p[hr,1-hr]} \leqq$$

$$\leqq 2^r\|f-S_{n,r}(f)\|_{L_p\left[\frac{hr}{2},1-\frac{hr}{2}\right]}+(2^r-1)\sup_{1\leqq j\leqq r}\|\Delta_h^r S_n^j(f)\|_{L_p[hr,1-hr]}.$$

Using Lemmas 4.4 and 4.5 with $a=rh_n$, $b=1-rh_n$ and A as chosen in those lemmas, we restrict h_n by $h_n\leqq\dfrac{hr}{(A+r)2}$, and as

$$[(r+A)h_n,\ 1-(r+A)h_n] \supseteq \left[\frac{hr}{2},\ 1-\frac{hr}{2}\right],$$

we have

$$\|f-S_{n,r}(f)\|_{L_p\left[\frac{hr}{2},1-\frac{hr}{2}\right]} \leqq M\sup_{0<h\leqq h_n}\|\Delta_h^r f\|_{L_p[h_nr,1-h_nr]}.$$

To estimate $\Delta_h^r S_n^j(f)$ we recall from the proof of Lemma 4.5 that $S_n^j(f,t)$ depends on the values of f in $\left[t-j\dfrac{m+1}{2}h_n,\ t+j\dfrac{m+1}{2}h_n\right]\subseteq\left[t-\dfrac{r}{2}(m+1)h_n,\ t+\dfrac{r}{2}(m+1)h_n\right]$, and as $t \in \left[\frac{hr}{2}, 1-\frac{hr}{2}\right]$, we have to choose $h_n \leqq \frac{r}{r+2}(m+1)h$ so that

$$\left[\frac{hr}{2}-\frac{r}{2}(m-1)h_n,\ 1-\frac{hr}{2}+\frac{r}{2}(m+1)h_n\right] \subset [h_n, 1-h_n] \subset [0, 1].$$

In this case the considerations in the proof of (5.1) in Theorem 5.1 are valid and

$$\|\Delta_h^r S_n^j(f)\|_{L_p[hr, 1-hr]} \leqq C(h^r/h_n^r)\|\Delta^r a_k(n)\|_{l_p(D)} h_n^{1/p} \leqq C(h/h_n)^r \varphi(h_n)$$

which together with the above, implies (7.6) for

$$h_n \leqq \min\left(\frac{hr}{(A+r)2}, \frac{r}{r+2}(m+1)h\right) \leqq h_{n-1}.$$

Taking supremum on h and recalling the definition of $\Omega_r(f, t)_p$, we have for $h_n \leqq Bh < h_{n-1}$

$$\Omega_r(f, t)_p \leqq C\Omega_r(f, h_n)_p + C\left(\frac{h}{h_n}\right)^r \varphi(h_n). \tag{7.7}$$

In order to continue our proof, we may choose a subsequence of h_n which we still call h_n satisfying $h_{n+1}/h_n \leqq B$ or $h_n/h_{n+1} \geqq B^{-1}$ and continue as was done in the proof of Theorem 5.1. (The only difference is that T_1 chosen in the present proof is restricted further.) This leads to

$$\Omega_r(f, t)_p = O(\varphi(h_r)) \quad \text{for} \quad h_n \leqq Bt < h_{n-1},$$

and therefore to $\Omega_r(f, t) = O(\varphi(t))$. Using the restriction on $\varphi(t)$, we have

$$\int_0^t (\varphi(\tau)/\tau)\, d\tau \leqq c\varphi(t),$$

and therefore, $\omega_r(f, t)_p = O(\varphi(t))$.

To obtain the proof of (2.6), we follow the proof of (5.2) and show that $\Omega_r(f, t)_p \sim \varphi(t)$ implies $\|\Delta^r a_k(n)\|_{l_p(D)} \geqq M_1 \varphi(h_n) h_n^{-1/p}$ where (7.7) is used instead of (5.5).

8. Extension of results about $C(D)$ and discrete data

In an earlier paper [8] results relating moduli of smoothness and information on discrete data were established. The use of some of the techniques developed here will yield the somewhat more general result which is stated below.

THEOREM 8.1. *Suppose D is a finite or infinite interval, $f \in C(D)$, $\varphi \in \Phi$ as given in Theorem 2.1, $\xi_n \in R$, $h_n = o(1)$ and $1 \leqq h_n/h_{n+1} \leqq M$, and suppose one of the following is satisfied:*

(a) $\omega(f, t)_{L_\infty(D)} = O(t^\delta)$ *for some* $\delta > 0$;

(b) $f \in$ A.C. (D);

(c) *f is sectionally monotonic, that is, in every finite proper subinterval of D there are only a finite number of changes of direction;*

(d) $\xi_n = \xi$ *and* $h_n = m(n)h_{n+1}$ *for some integer* $m(n)$ *satisfying* $m(n) \leqq M$.

Then

$$\text{(8.1)} \qquad \sup |\vec{\Delta}^r_{h_n} f(\xi_n + kh_n)| = O(\varphi(h_n)) \Leftrightarrow \omega_r(f, t)_\infty = O(\varphi(t))$$

and

$$\text{(8.2)} \qquad \sup |\vec{\Delta}^r_{h_n} f(\xi_n + kh_n)| \sim \varphi(h_n) \Leftrightarrow \omega_r(f, t)_\infty \sim \varphi(t)$$

where the supremum is taken on all k such that $[\xi_n + kh_n, \xi_n + (k+r)h_n] \subset D$.

REFERENCES

[1] Boman, J., On a problem concerning moduli of smoothness, *Fourier analysis and approximation theory* (Proc. Colloq., Budapest, Hungary, 1976), Vol. 1, Colloq. Math. Soc. János Bolyai, 19, North-Holland, Amsterdam, 1978, 175—179. *MR* **81b:** 42025.

[2] Boor, C. de, *A practical guide to splines,* Applied Mathematical Sciences, 27, Springer-Verlag, New York—Heidelberg, 1978. *MR* **80a:** 65027.

[3] Ciesielski, Z., Approximation by splines and its application to Lipschitz classes and to stochastic processes, *The theory of the approximation of functions* (Proc. Internat. Conf., Kaluga, 1975), Nauka, Moscow, 1977, 397—404. *MR* **81c:** 41019.

[4] Ciesielski, Z., Properties of realizations of random fields, *Mathematical statistics and probability theory* (Proc. Sixth Internat. Conf., Wisła, 1978), Lecture notes in statistics, Vol. 2, Springer-Verlag, New York, 1980, 97—110. *MR* **82g:** 60064.

[5] DeVore, R., *The approximation of continuous functions by positive linear operators,* Lecture Notes in Mathematics, Vol. 293, Springer-Verlag, Berlin—New York, 1972. *MR* **54** # 8100.

[6] Ditzian, Z., Inverse theorems for functions in L_p and other spaces, *Proc. Amer. Math. Soc.* **54** (1976), 80—82. *MR* **52** # 14765.

[7] Ditzian, Z., Discrete and shift Kolmogorov-type inequalities, *Proc. Roy. Soc. Edinburgh Sect. A* **93** (1982/83), 307—317. *MR* **84m:** 47038.

[8] Ditzian, Z., Moduli of smoothness using discrete data, *J. Approx. Theory* **49** (1987), 115—129.

[9] Ditzian, Z., On divided differences on $C[a, b]$, $L_p[a, b]$ and other spaces, *Period. Math. Hungar.* **17** (1986), 45—64.

[10] Ditzian, Z. and Totik, V., Moduli of smoothness, Springer-Verlag, 1987.

[11] Freud, G., On the problem of R. de Vore, *Canad. Math. Bull.* **17** (1974), 39—44. *MR* **50** # 10168.

[12] Schumaker, L., *Spline functions: basic theory,* Pure and Applied Mathematics. A Wiley-Interscience Publication, John Wiley & Sons, Inc., New York, 1981. *MR* **82j:** 41001.

[13] Totik, V., On a problem concerning L_p moduli of continuity, *Studia Sci. Math. Hungar.* **19** (1984), 97—101.

(Received May 25, 1986)

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF ALBERTA
EDMONTON, ALBERTA
T6G 2G1
CANADA

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 67—69

ЧИСЛОВАЯ ОБЛАСТЬ ЛИНЕЙНЫХ ОПЕРАТОРОВ В ПРОСТРАНСТВАХ С ИНДЕФИНИТНОЙ МЕТРИКОЙ

Ц. БАЯСГАЛАН

В настоящей заметке доказана выпуклость числовой области произвольного линейного оператора в пространстве с индефинитной метрикой. Как ее приложение, приведена одна простая связь между спектром и числовой областью положительного оператора в пространстве Крейна (см. [1]).

В комплексном линейном пространстве H заданы эрмитова билинейная форма (x, y), определенная для всех $x, y \in H$, и линейный оператор A, определенный всюду в нем.

Теорема. *Множество*

$$V(A) = \{(Ax, x)\colon (x, x) = 1\}$$

выпукло.

Доказательство. Пусть существуют векторы $x, y \in H$, такие что $(x, x) = (y, y) = 1$, $(Ax, x) > 0$, $(Ay, y) < 0$. Тогда найдется $z \in H$ со свойствами $(z, z) = 1$, $(Az, z) = 0$.

Действительно, мы ищем z в виде $z = t_1 x + t_2 y$, где t_1 и t_2 отличны от нуля. Введем следующие обозначения:

$$(Ax, x) = a_{11}, \quad (Ay, y) = a_{22}, \quad (Ax, y) = a_{12}, \quad (Ay, x) = a_{21}.$$

Тогда уравнение $(Az, z) = 0$ можно переписать в виде

$$a_{11} + \frac{t_2}{t_1} a_{21} + \left(\overline{\frac{t_2}{t_1}}\right)\left(\frac{t_2}{t_1} a_{22} + a_{12}\right) = 0.$$

Или, вводя обозначение $\frac{t_2}{t_1} = \theta$, получаем уравнение

$$a_{11} + \theta a_{21} + \bar{\theta}(\theta a_{22} + a_{12}) = 0$$

относительно θ. В обозначениях $a_{21} = c + di$, $a_{12} = \delta + \varepsilon i$, $\theta = \theta_1 + \theta_2 i$, последнее уравнение сводится к системе уравнений

$$(1) \qquad \begin{aligned} &a_{11} + a_{22}(\theta_1^2 + \theta_2^2) + \theta_1(c + \delta) + \theta_2(\varepsilon - d) = 0, \\ &\theta_1(\varepsilon + d) + \theta_2(c - \delta) = 0. \end{aligned}$$

1980 *Mathematics Subject Classification.* Primary 47A55; Secondary 47A60.
Key words and phrases. Krein space, nonnegative operators in Krein space.

Из условий $(z, z)=1$ и $\theta=\frac{t_2}{t_1}$ следует равенство

$$|t_1|^2\bigl(1+|\theta|^2+2\operatorname{Re}(\theta(y, x))\bigr) = 1.$$

Отсюда необходимо вытекает неравенство

$$1+|\theta|^2+2\operatorname{Re}\bigl(\theta(y, x)\bigr) > 0. \tag{2}$$

Таким образом, нам достаточно найти такой θ, что он удовлетворяет (1) и (2). Если $\varepsilon+d=0$, то положим $\theta_2=0$, и в зависимости от знака $\operatorname{Re}(y, x)$ мы выбираем в качестве θ_1 одно из решений квадратного уравнения из системы (1). Если же $\varepsilon+d\neq 0$, то (1) и (2) перепишем в виде

$$\begin{aligned} & a_{11}+a_{22}\left[1+\left(\frac{\delta-c}{\varepsilon+d}\right)^2\right]\theta_2^2+\theta_2(\varepsilon-d)+\frac{\delta^2-c^2}{\varepsilon+d}\,\theta_2=0,\\ & \theta_1=\frac{\delta-c}{\varepsilon+d}\,\theta_2,\\ & 1+|\theta|^2+2\theta_2\left[\frac{\delta-c}{\varepsilon+d}\,m-n\right]>0, \end{aligned} \tag{3}$$

где $(y, x)=m+ni$. Аналогично предыдущему, в зависимости от знака $\frac{\delta-c}{\varepsilon+d}\,m-n$, мы выбираем одно из решений квадратного уравнения из системы (3).

Далее, пусть векторы $x, y\in H$ обладают свойствами

$$(x, x) = (y, y) = 1, \quad (Ax, x) = 1, \quad (Ay, y) = -1.$$

Тогда для произвольного $t\in(-1, 1)$ мы имеем неравенства $\bigl((A-t)x, x\bigr)>0$, $\bigl((A-t)y, y\bigr)<0$. В силу доказанного, существует $z\in H$ такое, что $(z, z)=1$, $(Az, z)=t$.

Теперь, как в [2] (стр. 305), для произвольных $x, y\in H$ с $(Ax, x)\neq(Ay, y)$ и $(x, x)=(y, y)=1$ найдем постоянные α и β такие, что

$$\bigl((\alpha A+\beta)x, x\bigr) = 1, \quad \bigl((\alpha A+\beta)y, y\bigr) = -1.$$

Тогда, по предыдущему, для $t\in[-1, 1]$ существует $z\in H$ с $(z, z)=1$ и $\bigl((\alpha A+\beta)z, z\bigr)=t$. Подставляя значения α и β, мы приходим к равенству

$$\frac{1+t}{2}(Ax, x)+\frac{1-t}{2}(Ay, y) = (Az, z).$$

Следствие. *Пусть A — положительный оператор в пространстве Крейна. Тогда имеет место включение*

$$\sigma(A) \subset \overline{V_+(A)}\cup\overline{V_-(A)},$$

где $\sigma(A)$ — спектр оператора A и

$$V_+(A) = \{(Ax, x)\colon (x, x) = 1\}, \quad V_-(A) = \{-(Ax, x)\colon (x, x) = -1\}.$$

Заметим, что в силу предыдущей теоремы множества $V_+(A)$ и $V_-(A)$ выпуклы.

Доказательство. Введем обозначения

$$\alpha_+ = \inf_{(x,x)=1}(Ax, x), \qquad \alpha_- = \sup_{(x,x)=-1}(-(Ax, x)).$$

Если $\alpha_- = \alpha_+$, то по доказанной теореме множество $\overline{V_+(A)} \cup \overline{V_-(A)}$ выпукло, следовательно по [3] имеем

$$\sigma(A) \subset [\min \sigma(A), \max \sigma(A)] \subset \overline{V_-(A)} \cup \overline{V_+(A)}.$$

Если же $\alpha_- < \alpha_+$, то оператор A фундаментально приводим (см. [4], стр. 146), в силу чего следствие верно и в этом случае.

ЛИТЕРАТУРА

[1] BOGNÁR, J., *Indefinite inner product spaces*, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 78, Springer-Verlag, New York—Heidelberg, 1974. *MR* 57 # 7125.

[2] Халмош, П., Гильбертово пространство в задачах, Мир, Москва, 1970. *MR* 42 # 3586.

[3] TEXTORIUS, B., Minimaxprinzipe zur Bestimmung der Eigenwerte *J*-nichtnegativer Operatoren, *Math. Scand.* **35** (1974), 105—114. *MR* **51** # 6486.

[4] Баясгалан, Ц., О фундаментальной приводимости положительных операторов в пространствах с индефинитной метрикой, *Studia Sci. Math. Hungar.* **13** (1978), 143—150.

(Поступила 15-ого мая, 1986 г.)

МОНГОЛЬСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
П/Я 46—880
УЛАН БАТОР
MONGOLIA

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 71—91

ON TRIGONOMETRIC POLYNOMIALS WITH POSITIVE COEFFICIENTS

T. ERDÉLYI and J. SZABADOS

1. Introduction

In [2], G. G. Lorentz introduced the polynomials

$$\sum_{\substack{k+l\leq n\\ k,l\geq 0}} a_{k,l}(1-x)^k(1+x)^l, \quad \text{all} \quad a_{k,l}\geqq 0,$$

and proved that they have some interesting approximation theoretic properties. The structure of these polynomials was investigated in [1]. The aim of the present paper is to show that in case of trigonometric polynomials, similar investigations on intervals shorter than the period, can be carried out. Although the results are more or less similar to the case of algebraic polynomials, the proofs are more involved and require different ideas.

Let us start with some definitions. $\mathcal{T}_n$ will denote the set of trigonometric polynomials (with real coefficients) of order at most n. Let

$$0<\omega\leqq\pi;$$

in what follows we will work on the interval

(1) $$I_\omega=[-\omega,\omega].$$

(When speaking of intervals of length not greater than 2π, it is obviously sufficient to consider intervals of the form (1).) Now if $p(t)\in\mathcal{T}_n$, $\omega<\pi$ and $m\geqq n$ then $p(x)$ can be *uniquely* represented in the form

(2) $$p(t)=\sum_{k=0}^{2m} a_k \sin^k\frac{\omega-t}{2}\sin^{2m-k}\frac{\omega+t}{2}$$

with some a_k's. To see this, let

(3) $$u=\sin\frac{\omega-t}{2}\Big/\sin\frac{\omega+t}{2},$$

then

$$\sin t=\frac{(1-u^2)\sin\omega}{u^2+2u\cos\omega+1},\quad \cos t=\frac{(1+u^2)\cos\omega+2u}{u^2+2u\cos\omega+1},$$

$$\sin^2\frac{\omega+t}{2}=\frac{\sin^2\omega}{u^2+2u\cos\omega+1}.$$

Research supported by Hungarian National Foundation for Scientific Research Grant No. 1801.

1980 *Mathematics Subject Classification* (1985 Revision). Primary 42A05, 26C10. Secondary 26D05.

Key words and phrases. Positive coefficients, roots, degree of trigonometric polvnomials.

Thus substituting (3) in (2) and multiplying through by $(u^2+2u\cos\omega+1)^m$ we obtain

$$\sin^{2m}\omega \sum_{k=0}^{2m} a_k u^k = P(u)$$

where $P(u)$ is an algebraic polynomial of degree at most $2m$. Hence the a_k's are uniquely determined.

The set of those trigonometric polynomials which can be represented in the form (2) with some m and with $a_k \geqq 0$ $(k=0, \ldots, 2m)$ or $a_k \leqq 0$ $(k=0, \ldots, 2m)$ will be denoted by L_ω^+ or L_ω^-, resp. The set

$$L_\omega = L_\omega^+ \cup L_\omega^-$$

will be called trigonometric Lorentz polynomials (on I_ω). Note that L_π contains only the polynomials $c\cdot\cos^{2n}\frac{t}{2}$ $(n=0, 1, \ldots)$. To each $p(t)\in L_\omega$ there exists a minimal m such that in the representation (2) we have either all $a_k\geqq 0$ or all $a_k\leqq 0$. This will be called the Lorentz order of $p(t)$ and it will be denoted by $\sigma(p)$, in contrast with the usual order ord p. Evidently,

$$\operatorname{ord} p \leqq \sigma(p).$$

2. First order Lorentz polynomials

At first we would like to give a necessary and sufficient condition for a $p\in\mathcal{T}_1$ to have $\sigma(p)=1$. For this purpose we introduce the domains

(4) $$D_\omega = \{z = x+iy \mid \cos\omega \operatorname{ch} y \geqq \cos x, \ -\pi \leqq x < \pi\}$$

in the complex plane. Depending on ω, these domains may be of different character; see the shaded areas on Fig. 1.

THEOREM 1. *For a polynomial $p(t)\in\mathcal{T}_1\setminus\mathcal{T}_0$ we have $\sigma(p)=1$ if and only if the roots of $p(t)$ are in D_ω.*

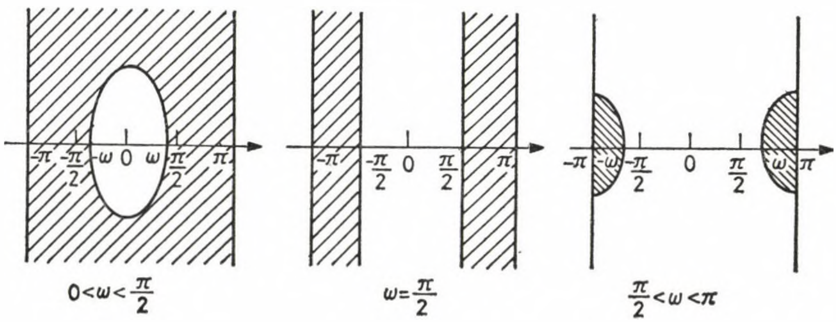

Fig. 1

PROOF. Observe that the "half order" factors of $p(t)$

$$\sin\frac{t-x}{2}=\frac{1}{\sin\omega}\left(\sin\frac{-\omega-x}{2}\sin\frac{\omega-t}{2}+\sin\frac{\omega-x}{2}\sin\frac{\omega+t}{2}\right)$$

containing the real root x with $-\pi\leqq x\leqq-\omega$ or $\omega\leqq x<\pi$ we have $\operatorname{sgn}\sin\frac{-\omega-x}{2}=$ $=\operatorname{sgn}\sin\frac{\omega-x}{2}$.

Furthermore it is trivial that $\sigma(p)=1$ implies that p has no roots in $(-\omega,\omega)$. For the complex roots of $p(t)$ the statement follows from the obvious representation

$$p(t)=\frac{c}{\sin^2\omega}\left\{[\operatorname{ch}y-\cos(\omega-x)]\sin^2\frac{\omega-t}{2}+\right.$$

$$+2[\cos\omega\operatorname{ch}y-\cos x]\sin\frac{\omega-t}{2}\sin\frac{\omega+t}{2}+ \tag{5}$$

$$\left.+[\operatorname{ch}y-\cos(\omega+x)]\sin^2\frac{\omega+t}{2}\right\}$$

where $x\pm iy$ are the roots of $p(t)$ and c is a constant. □

Let

$$D=\{z=x+iy\,|-\pi\leqq x<\pi\}.$$

The next problem which naturally arises: what happens if the roots of a first order polynomial are not in D_ω? The answer is different for different ω's. We use the same notation for the roots of $p(t)$ as above.

THEOREM 2. (a) *If* $0<\omega<\frac{\pi}{2}$ *and the roots of* $p(t)$ *are in* $D\setminus(D_\omega\cup\operatorname{int}I_\omega)$ *then* $p\in L_\omega$ *and*

$$\sigma(p)<\max\frac{4\sin(\omega\pm x)(\sin\omega\operatorname{ch}y\mp\sin x)}{\cos\omega\operatorname{sh}^2y}-1 \tag{6}$$

where the maximum is to be taken both for the upper and the lower signs.

(b) *If* $\frac{\pi}{2}\leqq\omega\leqq\pi$ *and the roots of* $p(t)$ *are in* $D\setminus D_\omega$ *then* $p\notin L_\omega$.

PROOF. (a) Consider the first order representation (5) of $p(t)$. Here the middle coefficient $\cos\omega\operatorname{ch}y-\cos x$ is possibly negative. In order to get higher order representations of $p(t)$, we multiply (5) by

$$1=\left[\frac{1}{\sin^2\omega}\left(\sin^2\frac{\omega-t}{2}+2\cos\omega\sin\frac{\omega-t}{2}\sin\frac{\omega+t}{2}+\sin^2\frac{\omega+t}{2}\right)\right]^m. \tag{7}$$

We would like to choose m so large that the coefficients of the resulting representation of $p(t)$ are of the same sign. Using the binomial theorem, the product of (5) and (7) can be written in the form

$$\begin{aligned}p(t) = p(t)\sin^{-2m}\omega \sum_{k=0}^{m}\binom{m}{k}\sin^k\frac{\omega-t}{2}\sin^{m-k}\frac{\omega+t}{2}\times\\ \times\left(\sin\frac{\omega-t}{2}+\cos\omega\sin\frac{\omega+t}{2}\right)^k\left(\cos\omega\sin\frac{\omega-t}{2}+\sin\frac{\omega+t}{2}\right)^{m-k}.\end{aligned}\tag{8}$$

We will perform the multiplication by $p(t)$ separately for each term in the sum for k. Let first $\frac{m}{2}\le k\le m$; then the corresponding terms will be (apart from factors with positive coefficients)

$$p(t)\left(\sin\frac{\omega-t}{2}+\cos\omega\sin\frac{\omega+t}{2}\right)^k=$$

$$\begin{aligned}=\frac{c}{\sin^2\omega}\left\{[\operatorname{ch} y-\cos(\omega-x)]\sin^2\frac{\omega-t}{2}+2\left(\operatorname{ch} y-\frac{\cos x}{\cos\omega}\right)\sin\frac{\omega-t}{2}\times\right.\\ \left.\times\left(\cos\omega\sin\frac{\omega+t}{2}\right)+\frac{\operatorname{ch} y-\cos(\omega+x)}{\cos^2\omega}\left(\cos\omega\sin\frac{\omega+t}{2}\right)^2\right\}\times\\ \times\left(\sin\frac{\omega-t}{2}+\cos\omega\sin\frac{\omega+t}{2}\right)^k.\end{aligned}\tag{9}$$

Replace here $\sin\frac{\omega-t}{2}$ and $\cos\omega\sin\frac{\omega+t}{2}$ by $\frac{1-v}{2}$ and $\frac{1+v}{2}$, resp. The resulting algebraic polynomial (apart from a multiplicative constant)

$$\begin{aligned}P_k(v)=\left[\frac{\operatorname{ch} y-\cos(\omega-x)}{4}(1-v)^2+\frac{\cos\omega\operatorname{ch} y-\cos x}{2\cos\omega}(1-v^2)+\right.\\ \left.+\frac{\operatorname{ch} y-\cos(\omega+x)}{4\cos^2\omega}(1+v)^2\right]\left(\frac{1-v}{2}+\frac{1+v}{2}\right)^k\end{aligned}\tag{10}$$

will have nonnegative coefficients in $1-v$ and $1+v$ if and only if (9) has nonnegative coefficients in $\sin\frac{\omega-t}{2}$ and $\sin\frac{\omega+t}{2}$. Hence the problem is: how large k should be in (10) so that the coefficients are nonnegative? In other words, we are interested in the algebraic Lorentz degree of the quadratic polynomial $P_k(v)$. An upper estimate for this degree is

$$d(P_k)<\frac{2(1-a^2)}{b^2}+1 \quad\text{if}\quad a^2+b^2<1,\quad b\ne 0,\tag{11}$$

where $a\pm bi$ are the roots of $P_k(v)$ (see [1, Theorem 2(ii)]). An easy calculation shows that $b\neq 0$, $a^2+b^2<1$ is equivalent to $a\pm bi\notin D_\omega\cup\operatorname{int} I_\omega$, thus in our case (11) will be

$$\frac{m}{2}+2\leqq k+2=d(P_k)<\frac{2\sin(\omega+x)(\sin\omega\operatorname{ch}y-\sin x)}{\cos\omega\operatorname{sh}^2 y}+1,$$

i.e.

$$m+1<\frac{4\sin(\omega+x)(\sin\omega\operatorname{ch}y-\sin x)}{\cos\omega\operatorname{sh}^2 y}-1. \tag{12}$$

Similarly, to have nonnegative coefficients in (10) for $0\leqq k<\frac{m}{2}$ we get the condition

$$m+1<\frac{4\sin(\omega-x)(\sin\omega\operatorname{ch}y+\sin x)}{\cos\omega\operatorname{sh}^2 y}-1.$$

This together with (12) proves (6).

(b) In order to get higher order representations of (5), we have to multiply by

$$1=\frac{1}{\sin^2\omega}\left(\sin^2\frac{\omega-t}{2}+2\cos\omega\sin\frac{\omega-t}{2}\sin\frac{\omega+t}{2}+\sin^2\frac{\omega+t}{2}\right). \tag{13}$$

If we assume $c>0$ in (5), then the signs of the coefficients of $p(t)$ are $+-+$. Because of $\frac{\pi}{2}\leqq\omega<\pi$, the signs in (13) are also $+-+$ (considering a zero coefficient as $-$). Thus multiplying by (13) yields a sequence of coefficients with signs $+-+-+$. Multiplying successively by (13), the resulting coefficients will be $+-+\ldots-+$, i.e. never $++\ldots+$ or $--\ldots--$. Hence $p(t)\notin L_\omega$. □

We do not deal with the lower estimate of the Lorentz degree of first order polynomials here. This will be settled in the next section in general for polynomials of order n.

3. Higher order Lorentz polynomials

In order to formulate our results about higher order Lorentz polynomials we introduce some definitions. Let $\varphi(x)$ be a positive continuous function in int I_ω and denote

$$D(\varphi)=\{z=x+iy\,|\,|y|\geqq\varphi(x)\quad\text{if}\quad |x|<\omega;$$
$$y\text{ arbitrary if }-\pi\leqq x\leqq-\omega\text{ or }\omega\leqq x<\pi\}.$$

$\left(\text{Note that for } \varphi(x)=\operatorname{arch}\frac{\cos x}{\cos\omega}\ \left(\omega<\frac{\pi}{2}\right) \text{ we have } D(\varphi)=D_\omega.\right)$ Also let

$$L_n(\varphi)=\{p(t)\,|\,p\in\mathcal{T}_n,\ p(z)\neq 0\ \text{if}\ z\in D\setminus D(\varphi)\}$$

and

$$\sigma_n(\varphi) = \sup_{p \in L_n(\varphi)} \sigma(p).$$

THEOREM 3. *Let* $0 < \omega < \frac{\pi}{2}$. *Then*

$$\sigma_n(\varphi) \leqq n\left(\frac{4}{\cos\omega} \sup_{|a|<\omega} \frac{(\omega^2-a^2)^2}{\varphi(a)^2} + 2\tan\omega + 1\right). \tag{14}$$

On the other hand we have *

(a) $$\sigma_n(\varphi) \geqq c_1 n \sup_{|a|<\omega} \frac{(\omega^2-a^2)^2}{\omega^2\varphi(a)^2},$$

(b) $$\sigma_n(\varphi) \geqq c_2 n \sup_{|a|<\omega} \frac{\frac{1}{\omega}\tan\omega(\omega^2-a^2)}{\operatorname{sh}^2\varphi(a) + (\omega^2-a^2)^2 \frac{1}{\omega^2}},$$

and

(c) $$\sigma_n(\varphi) \geqq c_3 n \sup_{|a|<\omega} \frac{\omega^2-a^2}{\operatorname{sh}\varphi(a)\left(\operatorname{sh}\varphi(a) + \frac{\omega^2-a^2}{\omega}\right)}.$$

PROOF. First we prove the upper estimate. Evidently, each $p(t) \in L_n(\varphi)$ can be written in the form

$$p(t) = c \prod_{j=1}^{2N} \sin\frac{t-x_j}{2} \prod_{j=2N+1}^{n} p_j(t),$$

where either $-\pi \leqq x_j \leqq -\omega$ or $\omega \leqq x_j < \pi$ $(j=1, \dots, 2N)$, and $p_j(t)$ has complex roots $x_j \pm iy_j$, $y_j \neq 0$ $(j=2N+1, \dots, n)$. Now

$$\sin\frac{t-x_j}{2} = \frac{1}{\sin\omega}\left(\sin\frac{-\omega-x_j}{2}\sin\frac{\omega-t}{2} + \sin\frac{\omega-x_j}{2}\sin\frac{\omega+t}{2}\right) \tag{15}$$

and here $\operatorname{sgn}\sin\frac{-\omega-x_j}{2} = \operatorname{sgn}\sin\frac{\omega-x_j}{2}$. As for $p_j(t)$, applying Theorem 2(a) we obtain

$$\sigma(p_j) < \max\frac{4\sin(\omega\pm x_j)\sin\omega\operatorname{ch}y_j \mp \sin x_j)}{\cos\omega\operatorname{sh}^2 y_j} + 1 =$$

$$= \max\left[\frac{4\sin(\omega\pm x_j)\sin\omega}{\cos\omega(\operatorname{ch}y_j+1)} + \frac{4\sin(\omega\pm x_j)(\sin\omega\mp\sin x_j)}{\cos\omega\operatorname{sh}^2 y_j}\right] + 1 \leqq$$

$$\leqq 2\tan\omega + 1 + \frac{4(\omega^2-x_j^2)}{\cos\omega\varphi(x_j)^2} \leqq \frac{4}{\cos\omega}\sup_{|a|<\omega}\frac{\omega^2-a^2}{\varphi(a)^2} + 2\tan\omega + 1$$

which together with (15) yields (14).

---

* In what follows, $c_1, c_2, \dots$ will denote absolute positive constants.

To prove (a), let

$$p(t)=[\operatorname{ch}\varphi(a)-\cos(t-a)]^n\in L_n(\varphi) \tag{16}$$

where $a\in\operatorname{int} I_\omega$ is arbitrary. At first assume that

$$\varphi(a)\leqq\frac{\omega^2-a^2}{8\omega} \tag{17}$$

and let

$$p(t)=\sum_{k=0}^{2m}a_k\sin^k\frac{\omega-t}{2}\sin^{2m-k}\frac{\omega+t}{2}\quad(a_k\geqq 0,\ k=0,1,\dots,2m) \tag{18}$$

be the Lorentz representation of $p(t)$ (i.e. $m=\sigma(p)$). Using Cauchy—Schwarz inequality we obtain

$$p(a+\varphi(a))\,p(a+3\varphi(a))=\sum_{k=0}^{2m}a_k\sin^k\frac{\omega-a-\varphi(a)}{2}\times$$

$$\times\sin^{2m-k}\frac{\omega+a+\varphi(a)}{2}\sum_{k=0}^{2m}a_k\sin^k\frac{\omega-a-3\varphi(a)}{2}\sin^{2m-k}\frac{\omega+a+3\varphi(a)}{2}\geqq$$

$$\geqq\left\{\sum_{k=0}^{2m}a_k\left[\sin\frac{\omega-a-\varphi(a)}{2}\sin\frac{\omega-a-3\varphi(a)}{2}\right]^{k/2}\times\right.$$

$$\left.\times\left[\sin\frac{\omega+a+\varphi(a)}{2}\sin\frac{\omega+a+3\varphi(a)}{2}\right]^{m-k/2}\right\}^2=$$

(19)

$$=\left\{\sum_{k=0}^{2m}a_k\sin^k\frac{\omega-a-3\varphi(a)}{2}\sin^{2m-k}\frac{\omega+a+2\varphi(a)}{2}\times\right.$$

$$\times\left[\frac{\sin\frac{\omega-a-\varphi(a)}{2}\sin\frac{\omega-a-3\varphi(a)}{2}}{\sin^2\frac{\omega-a-2\varphi(a)}{2}}\right]^{k/2}\times$$

$$\left.\times\left[\frac{\sin\frac{\omega+a+\varphi(a)}{2}\sin\frac{\omega+a+3\varphi(a)}{2}}{\sin^2\frac{\omega+a+2\varphi(a)}{2}}\right]^{m-k/2}\right\}^2.$$

Here by $\frac{2}{\pi}x<\sin x<x$ $(0<x<\pi/2)$, $\varphi(a)\leqq\min\frac{\omega\pm a}{4}<\frac{\pi}{8}$ (see (17)) and $1-x>$

$>e^{-2x}$ $(0<x<0.7)$ we get

$$\frac{\sin\frac{\omega-a-\varphi(a)}{2}\sin\frac{\omega-a-3\varphi(a)}{2}}{\sin^2\frac{\omega-a-2\varphi(a)}{2}}=1-\frac{\sin^2\frac{\varphi(a)}{2}}{\sin^2\frac{\omega-a-2\varphi(a)}{2}}>$$

$$>1-\frac{\varphi(a)^2}{4\sin^2\frac{\omega-a}{4}}>1-\frac{\pi^2\varphi(a)^2}{(\omega-a)^2}\geqq 1-\frac{4\pi^2\omega^2\varphi(a)^2}{(\omega^2-a^2)^2}>e^{-\frac{8\pi^2\omega^2\varphi(a)^2}{(\omega^2-a^2)^2}}$$

and similarly

$$\frac{\sin\frac{\omega+a+\varphi(a)}{2}\sin\frac{\omega+a+3\varphi(a)}{2}}{\sin^2\frac{\omega+a+2\varphi(a)}{2}}>e^{-\frac{8\pi^2\omega^2\varphi(a)^2}{(\omega^2-a^2)^2}}.$$

Thus (19) yields

$$p(a+\varphi(a))p(a+3\varphi(a))\geqq p(a+2\varphi(a))^2e^{-\frac{8\pi^2\omega^2\varphi(a)^2 2m}{(\omega^2-a^2)^2}}. \tag{20}$$

On the other hand, by (16), $\frac{x^2}{2}<\operatorname{ch}x-1<0.54x^2$ $\left(|x|<\frac{\pi}{8}\right)$ and $x^2\geqq\sin x^2>$ $>x^2-\frac{x^4}{3}$,

$$\frac{p(a+\varphi(a))(p+3\varphi(a))}{p(a+2\varphi(a))^2}=\left\{\frac{[\operatorname{ch}\varphi(a)-\cos\varphi(a)]\,[\operatorname{ch}\varphi(a)-\cos 3\varphi(a)]}{[\operatorname{ch}\varphi(a)-\cos 2\varphi(a)]^2}\right\}^n\leqq$$

$$\leqq\left\{\frac{\left[0.54\varphi(a)^2+2\sin^2\frac{\varphi(a)}{2}\right]\left[0.54\varphi(a)^2+2\sin^2\frac{3\varphi(a)}{2}\right]}{[0.5\varphi(a)^2+2\sin^2\varphi(a)]^2}\right\}^n\leqq$$

$$\leqq\left[\frac{1.04\cdot 5.04}{\left(2.5-\frac{2}{3}\varphi(a)^2\right)^2}\right]^n\leqq\left(\frac{1.04\cdot 5.04}{2.397^2}\right)^n<e^{-0.09n}.$$

This compared with (20) yields

$$\sigma(p)=m\geqq\frac{n(\omega^2-a^2)^2}{180\pi^2\varphi(a)^2\omega^2}. \tag{21}$$

Now if (17) does not hold then

$$\sigma(p)\geqq n\geqq\frac{n(\omega^2-a^2)^2}{64\omega^2\varphi(a)^2}$$

and thus (21) holds for any $|a|\leqq\omega$.

To prove (b) and (c) let $0 \leqq a < \omega$ and choose a b for which $a < b \leqq \omega$. We have

$$\sin\frac{\omega-t}{2} = \frac{1}{\sin\frac{\omega+b}{2}}\left(\sin\omega\sin\frac{b-t}{2} + \sin\frac{\omega-b}{2}\sin\frac{\omega+t}{2}\right)$$

where $\operatorname{sgn}\sin\omega = \operatorname{sgn}\sin\frac{\omega-b}{2}$. From this and the representation (18) easily follows that

$$p(x) = \sum_{k=0}^{2m} A_k \sin^k\frac{b-t}{2}\sin^{2m-k}\frac{\omega+t}{2} \qquad (A_k \geqq 0,\ k = 0, 1, \ldots, 2m).$$

Since

$$p(b) = A_0 \sin^{2m}\frac{\omega+b}{2}$$

and

$$p'(b) = A_0 2m \sin^{2m-1}\frac{\omega+b}{2}\,\frac{1}{2}\cos\frac{\omega+b}{2} - A_1\frac{1}{2}\sin^{2m-1}\frac{\omega+b}{2},$$

we obtain from (16) and $\omega < \frac{\pi}{2}$

$$\sigma(p) = m \geqq \tan\frac{\omega+b}{2}\,\frac{p'(p)}{p(b)} = \tan\frac{\omega+b}{2}\,n\,\frac{\sin(b-a)}{\operatorname{ch}\varphi(a) - \cos(b-a)} =$$

(22)

$$= \tan\frac{\omega+b}{2}\,n\,\frac{\sin(b-a)}{\operatorname{sh}^2\varphi(a) + \sin^2\frac{b-a}{2}}.$$

Choosing $b = \omega$, from (22) we get

$$\sigma(p) = m \geqq c_2 n \frac{\frac{1}{\omega}\tan\omega(\omega^2 - a^2)}{\operatorname{sh}^2\varphi(a) + (\omega^2 - a^2)^2\frac{1}{\omega^2}}$$

provided $0 \leqq a < \omega$. The same argument with $-\omega$ instead of ω leads to the same estimate for $-\omega < a \leqq 0$. Thus (b) is also proved.

Now we prove (c). Let $0 \leqq a < \omega$. We distinguish two cases.

Case 1. $a + \operatorname{sh}\varphi(a) \leqq \omega$. Then choosing $b = a + \varphi(a) < a + \operatorname{sh}\varphi(a) \leqq \omega$, from (22) we get

$$\sigma(p) = m \geqq c_4\omega n\frac{\varphi(a)}{\operatorname{sh}^2\varphi(a) + \varphi(a)^2} \geqq c_5\omega n\frac{1}{\operatorname{sh}\varphi(a)} \geqq$$

$$\geqq c_5 n\omega\frac{\frac{\omega^2 - a^2}{\omega}}{\operatorname{sh}\varphi(a)\left(\operatorname{sh}\varphi(a) + \frac{\omega^2 - a^2}{\omega}\right)}.$$

Case 2. $a+\operatorname{sh}\varphi(a)>\omega$. Now choosing $b=\omega$, from (22) we deduce

$$\sigma(p)=m\geqq c_6 n\omega\frac{\omega-a}{\operatorname{sh}\varphi(a)(\operatorname{sh}\varphi(a)+\omega-a)}\geqq$$

$$\geqq c_7 n\frac{\omega^2-a^2}{\operatorname{sh}\omega(a)\left(\operatorname{sh}\omega(a)+\dfrac{\omega^2-a^2}{\omega}\right)}.$$

The same argument with $-\omega$ instead of ω leads to the same estimate for $-\omega<a\leqq\omega$. □

COROLLARY 1. (a) *If* $0<\omega<\frac{\pi}{2}$ *and* $\sup_{|a|<\omega}\frac{\omega^2-a^2}{\varphi(a)^2}<\infty$ *then* $\sigma_n(\varphi)<\infty$ $(n=1,2,\ldots)$.

(b) *If* $0<\omega<\frac{\pi}{2}$ *and*

$$\sup_{|a|<\omega}\frac{\omega^2-a^2}{\varphi(a)^2}=\infty \tag{23}$$

then $\sigma_n(\varphi)=\infty$ $(n=1,2,\ldots)$.

PROOF. Only (b) needs an explanation. (23) implies $\sup_{|a|<\omega}\varphi(a)<\infty$, and thus from lower estimate (b) of Theorem 3

$$\sigma_n(\varphi)\geqq c_2 n\sup_{|a|<\omega}\frac{\frac{1}{\omega}\tan\omega(\omega^2-a^2)}{\operatorname{sh}^2\varphi(a)+(\omega^2-a^2)^2\frac{1}{\omega^2}}\geqq$$

$$\geqq c(\varphi,\omega)\,n\frac{\omega^2-a^2}{\varphi(a)^2+(\omega^2-a^2)^2}=c(\varphi,\omega)\,n\frac{1}{\dfrac{\varphi(a)^2}{\varphi^2-a^2}+\omega^2-a^2}\to\infty$$

as $a\to\omega$ or $a\to-\omega$ ($c(\varphi,\omega)$ is a positive constant depending only on φ and ω). □

CONJECTURE. We have *

$$c_8(\omega)\,n\sup_{|a|<\omega}\frac{\omega^2-a^2}{\varphi(a)^2}\leqq\sigma_n(\varphi)\leqq c_9(\omega)\,n\left(\sup_{|a|<\omega}\frac{\omega^2-a^2}{\varphi(a)^2}+1\right)\quad\left(0<\omega<\frac{\pi}{2}\right),$$

further if $\omega<c_{10}\frac{\pi}{2}$, where $c_{10}<1$, then

$$c_{11}n\sup_{|a|<\omega}\frac{\omega^2-a^2}{\varphi(a^2)}\leqq\sigma_n(\varphi)\leqq c_{12}n\left(\sup_{|a|<\omega}\frac{\omega^2-a^2}{\varphi(a)^2}+1\right).$$

* $c_i(\omega)$ $(i=1,2,\ldots)$ denote positive constants possibly depending on ω.

Several examples indicate that this conjecture is true. If e.g.

$$\varphi_1(x) = \varepsilon(\omega^2 - x^2)^\alpha \qquad \left(-\infty < \alpha \leqq \frac{1}{2},\ 0 < \varepsilon < 1,\ |x| < \omega\right)$$

or

$$\varphi_2(x) = \varepsilon\sqrt{\omega^2 - x^2} + \omega^2 - x^2 \qquad (0 < \varepsilon < 1,\ |x| < \omega),$$

where $\omega < c_{10}\frac{\pi}{2}$ with $c_{10} < 1$, then according to the upper estimate and lower estimates (a) and (c) of Theorem 3 we have $\sigma_n(\varphi_i) \sim \frac{n}{\varepsilon^2}$ $(i=1, 2)$. However, if $\omega < c_{10}\frac{\pi}{2}$ with $c_{10} < 1$ and

$$\varphi_3(x) = \varepsilon^{2/3}\left(\omega - \sqrt{1 - \varepsilon^{2/3}}\,x\right) \qquad (0 < \varepsilon < 1,\ |x| < \omega),$$

then by the same arguments we have only

$$c_{13}\frac{n}{\varepsilon^{4/3}} \leqq \sigma_n(\varphi_3) \leqq c_{14}\frac{n}{\varepsilon^2}.$$

Remark. From Theorem 3 (a), (c) it is easy to see that in case of $0 < \omega < c_{10}\frac{\pi}{2}$, $c_{10} < 1$ we have

$$\sigma_n(\varphi) \geqq c_{15}\, n\left(\sup_{|a|<\omega} \frac{\sqrt{\omega^2 - a^2}}{\operatorname{sh}\varphi(a)}\right)^{4/3}$$

for an arbitrary φ.

Corollary 2. *In case* $0 < \omega < \frac{\pi}{2}$, *a trigonometric polynomial is a Lorentz polynomial on* I_ω *if it has no roots in* $\operatorname{int} I_\omega$.

Proof. Only the sufficiency needs some explanation. If $p(t) \in \mathcal{T}_n$ has no roots in $\operatorname{int} I_\omega$, then, for sufficiently small $\varepsilon > 0$, $p(t) \in L(\varphi_0)$ where $\varphi_0(x) \equiv \varepsilon$. Since in this case

$$\sup_{|a|<\omega} \frac{\omega^2 - a^2}{\varphi_0(a)^2} = \frac{\omega^2}{\varepsilon^2},$$

Theorem 3 implies

$$\sigma(p) \leqq n\left(\frac{4\omega^2}{\varepsilon^2 \cos\omega} + 2\tan\omega + 1\right). \qquad \square$$

We now turn to the case $\frac{\pi}{2} \leqq \omega < \pi$.

Theorem 4. *Let* $\frac{\pi}{2} \leqq \omega < \pi$ *and* $p(t) \in \mathcal{T}_n \setminus \mathcal{T}_{n-1}$.

(a) *If all the roots of* $p(t)$ *are in* D_ω *then* $p \in L_\omega$ *and* $\sigma(p) = n$.
(b) *If all the roots of* $p(t)$ *are in* $D \setminus D_\omega$ *then* $p \notin L_\omega$.

(c) *If $p(t)$ has roots both in D_ω and $D\setminus D_\omega$ then both of the above possibilities can occur.*

(d) *If $p(t)$ has no roots in I_ω then for $q(t):=\sin^{2d}\frac{\pi-t}{2}p(t)$ we have $q\in L_\omega$ and* $\operatorname{ord} q=\sigma(q)=n+d$ *provided that d is large enough.*

Remark. The statement of Theorem 4 (d) is true even in case of $0<\omega<\frac{\pi}{2}$.

Proof. (a) This is a direct consequence of Theorem 1.

(b) We may assume that $p(t)$ has no real roots; otherwise the statement is obvious. (The real roots would be in $\operatorname{int} I_\omega$.) Represent $p(t)$ as a product of first order polynomials. These can be represented in the form (5), i.e. with coefficients of signs $+-+$, say. Thus their product $p(t)$ will have coefficients of signs $+-+-\dots-+$. In trying to get a Lorentz representation, we multiply by (13), which, again, has coefficients of signs $+-+\dots-+$. Again, the product will be $+-+\dots-+$, thus never a Lorentz representation.

(c) Consider the polynomial

$$\sin^2\frac{\omega-t}{2}(\operatorname{ch} y-\cos t)=\frac{\operatorname{ch} y-1}{\sin^2\omega}\sin^4\frac{\omega-t}{2}-\frac{2(1-\cos\omega\operatorname{ch} y)}{\sin^2\omega}\times$$

$$\times\sin^3\frac{\omega-t}{2}\sin\frac{\omega+t}{2}+\frac{\operatorname{ch} y-1}{\sin^2\omega}\sin^2\frac{\omega-t}{2}\sin^2\frac{\omega+t}{2}\qquad (y\text{ arbitrary})$$

which has roots both in D_ω and $D\setminus D_\omega$ (namely, ω and $\pm iy$, resp.). The coefficients are $+-+00$, and hence for the same reason as in (b), it cannot be a Lorentz polynomial.

On the other hand, if we take the Lorentz polynomial

$$2\sin^2\frac{\omega}{2}(1+\cos t)(\operatorname{ch} y+\cos t)=(\operatorname{ch} y+\cos\omega)\sin^4\frac{\omega-t}{2}+$$

$$+(\operatorname{ch} y+\cos\omega+1)\sin^3\frac{\omega-t}{2}\sin\frac{\omega+t}{2}+2\sin^2\frac{\omega-t}{2}\sin^2\frac{\omega+t}{2}+$$

$$+(\operatorname{ch} y+\cos\omega+1)\sin\frac{\omega-t}{2}\sin^3\frac{\omega+t}{2}+(\operatorname{ch} y+\cos\omega)\sin^4\frac{\omega+t}{2},$$

we can see that it has roots $-\pi\in D_\omega$ and $-\pi+iy\in D\setminus D_\omega$ $\left(\text{if } \operatorname{ch} y\geqq-\frac{1}{\cos\omega}\right)$.

(d) It is sufficient to consider the case $n=1$. If p has two real roots (outside I_ω) then according to (a) $d=0$ is suitable. Now suppose that p has the form (5), that is for suitable a, b and c

$$p(t)=a\sin^2\frac{\omega-t}{2}+b\sin\frac{\omega-t}{2}\sin\frac{\omega+t}{2}+c\sin^2\frac{\omega+t}{2}.$$

As p has no roots in $\operatorname{int} I_\omega$, we can deduce that ay^2+by+c has no roots in $(0, \infty)$ and this implies that

$$P(x):=a(1-x)^2+b(1-x)(1+x)+c(1+x)^2$$

has no roots in $(-1, 1)$. Recalling the algebraic result (11), for sufficiently large d and the algebraic polynomial

$$Q(x):=P(x)\left(\frac{1-x}{2}+\frac{1+x}{2}\right)^{2d}$$

we have $d(Q)=2+2d$, and from this it is obvious that for sufficiently large d

$$q(t)=p(t)\left(\sin\frac{\omega-t}{2}+\sin\frac{\omega+t}{2}\right)^{2d}\in L_\omega$$

and $\sigma(q)=1+d$. Now observe that

$$\left(\sin\frac{\omega-t}{2}+\sin\frac{\omega+t}{2}\right)^2=4\sin^2\frac{\omega}{2}\sin^2\frac{\pi-t}{2},$$

therefore

$$q(t)=\left(4\sin^2\frac{\omega}{k}\right)^k p(t)\sin^{2k}\frac{\pi-t}{2}$$

and by this the proof is complete. □

REMARK. Of course, Theorem 4(c) does not give a complete description of the situation, but we were unable to obtain a more exact relation between the position of the roots and the property of being a Lorentz polynomial.

4. A criterion for $\sigma_n(\varphi)=n$

In this section we examine under what conditions on φ will $\sigma_n(\varphi)$ reach its minimum (which is obviously n). Note that Theorem 4(a) already gave a condition for an individual polynomial to have minimal Lorentz degree. Here we consider the problem for a class of functions.

THEOREM 5. *Let* $0<\omega<\dfrac{\pi}{2}$.

(a) *If* $D(\varphi)\subseteq D_\omega$ *then*

$$\sigma_n(\varphi)=n \tag{24}$$

for all $n=1, 2, \ldots$.

(b) *If* (24) *holds for some* $n\geqq 1$ *then* $D(\varphi)\subseteq D_\omega$.

PROOF. (a) This follows from Theorem 1 if we apply it to the elementary factors of a polynomial $p\in L_n(\varphi)$.

(b) Consider the polynomial (16) with $|a|<\omega$ and some $n\geqq 1$. By assumption, $\sigma(p)=n$, thus

$$p(t)=\sum_{k=0}^{2n} a_k \sin^k\frac{\omega-t}{2}\sin^{2n-k}\frac{\omega+t}{2} \qquad (a_k\geqq 0,\ k=0,1,\ldots,2n).$$

Hence

$$p(\omega)=a_0\sin^{2n}\omega\leqq\left(\frac{\sin\omega}{\sin\frac{\omega+t}{2}}\right)^{2n}\sum_{k=0}^{2n}a_k\sin^k\frac{\omega-t}{2}\sin^{2n-k}\frac{\omega+t}{2}=$$

(25)

$$=\left(\frac{\sin\omega}{\sin\frac{\omega+t}{2}}\right)^{2n}p(t) \qquad (t\in I_\omega).$$

On the other hand, we get from (16) for $2|a|-\omega<t<\omega$

$$\frac{p(\omega)}{p(t)}=\left[\frac{\operatorname{ch}\varphi(a)-\cos(\omega-a)}{\operatorname{ch}\varphi(a)-\cos(t-a)}\right]^n=$$

(26)

$$=\left[1+\frac{2\sin\frac{\omega-t}{2}\sin\left(\frac{\omega+t}{2}-a\right)}{\operatorname{ch}\varphi(a)-\cos(t-a)}\right]^n.$$

Now (25) and (26) yield

$$\frac{\sin\left(\frac{\omega+t}{2}-a\right)}{\operatorname{ch}\varphi(a)-\cos(t-a)}\leqq\frac{\left(\frac{\sin\omega}{\sin\frac{\omega+t}{2}}\right)^2-1}{2\sin\frac{\omega-t}{2}}=\frac{\sin\frac{3\omega+t}{2}}{2\sin^2\frac{\omega+t}{2}}.$$

Thus letting $t\to\omega-0$ we get

$$\frac{\sin(\omega-a)}{\operatorname{ch}\varphi(a)-\cos(\omega-a)}\leqq\frac{\cos\omega}{\sin\omega},$$

i.e. $\cos a\leqq\cos\omega\operatorname{ch}\varphi(a)$, which (see (4)) shows that $D(\varphi)\subseteq D_\omega$. □

5. Schur-type inequalities

For a continuous function $f(x)$ on the finite interval $[a,b]$, denote $\|f\|_{[a,b]}=$ $=\sup_{x\in[a,b]}|f(x)|$. A classical result of I. Schur [3] states that if $P(x)$ is an algebraic polynomial of degree at most n then

$$\|P(x)\|_{[-1,1]}\leqq(n+1)\left\|P(x)\sqrt{1-x^2}\right\|_{[-1,1]}.$$

In this section we prove the trigonometric analogue of this result, and we show that for Lorentz polynomials the corresponding inequality can be improved.

THEOREM 6. *We have*

$$A_n(\omega) =: \sup_{0 \not\equiv p \in \mathcal{T}_n} \frac{\|p(t)\|_{I_\omega}}{\left\| p(t) \sqrt{\dfrac{\cos t - \cos \omega}{2}} \right\|_{I_\omega}} = \frac{2n+1}{\sin \dfrac{\omega}{2}} \tag{27}$$

$$(0 < \omega \leqq \pi,\ n = 0, 1, \ldots)$$

and here the supremum is attained if and only if

$$p(t) = c \cdot \frac{\cos (2n+1) \arcsin \dfrac{\sin \frac{t}{2}}{\sin \frac{\omega}{2}}}{\sqrt{\cos t - \cos \omega}}.$$

PROOF. Define $2n+1$ different points

$$t_k = 2 \arcsin \left(\sin \frac{\omega}{2} \sin \frac{k\pi}{2n+1} \right) \qquad (|t_k| < \omega,\ k = 0, \pm 1, \ldots, \pm n). \tag{28}$$

We distinguish two cases in estimating $\|p(t)\|_{I_\omega}$.

Case 1. $|t| \leqq t_n$. Then by (28)

$$\begin{aligned} \frac{\cos t - \cos \omega}{2} &\geqq \frac{\cos t_n - \cos \omega}{2} = \sin^2 \frac{\omega}{2} - \sin^2 \frac{t_n}{2} = \\ &= \sin^2 \frac{\omega}{2} \cos^2 \frac{n\pi}{2n+1} = \sin^2 \frac{\omega}{2} \sin^2 \frac{\pi}{2(2n+1)} > \frac{\sin^2 \frac{\omega}{2}}{(2n+1)^2}, \end{aligned} \tag{29}$$

whence

$$\begin{aligned} \max_{|t| \leqq t_n} |p(t)| &< \frac{2n+1}{\sin \dfrac{\omega}{2}} \max_{|t| \leqq t_n} \left| p(t) \sqrt{\frac{\cos t - \cos \omega}{2}} \right| \leqq \\ &\leqq \frac{2n+1}{\sin \dfrac{\omega}{2}} \left\| p(t) \sqrt{\frac{\cos t - \cos \omega}{2}} \right\|_{I_\omega}. \end{aligned} \tag{30}$$

Case 2. $t_n < |t| \leqq \omega$. Consider the functions

$$D_{nk}(t) = \frac{(-1)^k \sqrt{\dfrac{\cos t_k - \cos \omega}{2}} \sin (2n+1) \arcsin \dfrac{\sin \frac{t}{2}}{\sin \frac{\omega}{2}}}{(2n+1) \cos \dfrac{t_k}{2} \sin \dfrac{t - t_k}{2}} \tag{31}$$

$$(k = 0, \pm 1, \ldots, \pm n).$$

At first we show that $D_{nk}(t)\in\mathcal{T}_n$ $(k=0, \pm 1, \dots, \pm n)$. Since $\sin(2n+1)u = \sum_{j=0}^{n} \lambda_j \sin^{2j+1} u$, we have

$$D_{nk}(t) = \mu_{nk} \frac{\sin\frac{t}{2} \sum_{j=0}^{n} \lambda_j \sin^{2j}\frac{t}{2}}{\sin\frac{t-t_k}{2}} = \frac{\sin\frac{t}{2}\bar{D}_{nk}(t)}{\sin\frac{t-t_k}{2}} \qquad (k=0, \pm 1, \dots, \pm n), \tag{32}$$

where $\bar{D}_{nk}(t)\in\mathcal{T}_n$ $(k=0, \pm 1, \dots, \pm n)$. Hence (by $t_0=0$) $D_{n0}(t)\in\mathcal{T}_n$. On the other hand, by (28), (31) and (32), $\bar{D}_{nk}(t_k)=0$ $(k=\pm 1, \dots, \pm n)$, and thus $D_{nk}(T)\in\mathcal{T}_n$ indeed, for all $|k|\leqq n$.

An easy calculation shows that $D_{nk}(t_j)=\delta_{kj}$ $(k, j=0, \pm 1, \dots, \pm n)$, whence

$$p(t) = \sum_{k=-n}^{n} p(t_k) D_{nk}(t). \tag{33}$$

Thus denoting

$$M = \max_{|k|\leqq n}\left| p(t_k)\sqrt{\frac{\cos t_k - \cos\omega}{2}}\right| \leqq \left\| p(t)\sqrt{\frac{\cos t - \cos\omega}{2}}\right\|_{I_\omega}, \tag{34}$$

we obtain from (33)

$$p(t) \leqq M \left| \frac{\sin(2n+1)\arcsin\dfrac{\sin\frac{t}{2}}{\sin\frac{\omega}{2}}}{2n+1} \sum_{k=-n}^{n} \frac{1}{\cos\frac{t_k}{2}\sin\frac{t-t_k}{2}}\right| =$$

$$= M \left| \frac{D_{n0}(t)}{\sin\frac{\omega}{2}} + \frac{4D_{n0}(t)\sin^2\frac{t}{2}}{\sin\frac{\omega}{2}} \sum_{k=1}^{n} \frac{1}{\cos t_k - \cos t}\right| =$$

$$= \frac{M}{\sin\frac{\omega}{2}} \left| D_{n0}(t) + \frac{2D'_{n0}(t)\sin\frac{t}{2}}{\cos\frac{t}{2}}\right| = \frac{2M}{\sin\frac{\omega}{2}\cos\frac{t}{2}} \left| \frac{d}{dt}\left[D_{n0}(t)\sin\frac{t}{2}\right]\right| = \tag{35}$$

$$= M \left| \frac{\cos(2n+1)\arcsin\dfrac{\sin\frac{t}{2}}{\sin\frac{\omega}{2}}}{\sqrt{\dfrac{\cos t - \cos\omega}{2}}}\right| \qquad (t_n < |t| < \omega).$$

Evidently, here

$$p^*(t) = \frac{\cos(2n+1)\arc\sin\dfrac{\sin\dfrac{t}{2}}{\sin\dfrac{\omega}{2}}}{\sqrt{\dfrac{\cos t-\cos\omega}{2}}}\in\mathcal{T}_n,$$

and because of $\operatorname{sgn} p^*(t_k)=(-1)^k$ $(k=0, \pm 1, \ldots, \pm n)$, $p^{*\prime}(t)$ has at least $2n-1$ roots in $(-t_n, t_n)$. Since $p^*(t)$ is an even trigonometric polynomial, the only remaining root of $p^{*\prime}(t)$ must be at $t=\pi$. Hence $p^*(t)$ is monotone on $[-\omega, -t_n]$ and on $[t_n, \omega]$. Being

$$|p^*(\pm t_n)| = \frac{1}{\sqrt{\dfrac{\cos t_n-\cos\omega}{2}}} < \frac{2n+1}{\sin\dfrac{\omega}{2}} = |p^*(\pm\omega)|$$

(see (28)), we get from (35) and (34)

$$\max_{t_n<|t|\leqq\omega}|p(t)| \leqq M|p^*(\pm\omega)| \leqq \frac{2n+1}{\sin\dfrac{\omega}{2}}\left\|p(t)\sqrt{\frac{\cos t-\cos\omega}{2}}\right\|_{I_\omega}.$$

This together with (30) shows that (27) holds true.

It is clear from (30) that $\|p^*(t)\|_{I_\omega}$ is attained either on $[-\omega, -t_n]$ or $[t_n, \omega]$. Also, (35) shows that the first inequality turns into equality if and only if

$$(-1)^k p(t_k)\sqrt{\frac{\cos t_k-\cos\omega}{2}} = \text{const.} \qquad (k=0, \pm 1, \ldots, \pm n).$$

Apart from a multiplicative constant, these $2n+1$ conditions uniquely determine $p(t)$, namely $p(t)=p^*(t)$. □

It is worth mentioning that Theorem 6 remains true for $\omega=\pi$: we have

$$\max_{-\infty<t<\infty}|p(t)| \leqq (2n+1)\max_{-\infty<t<\infty}\left|p(t)\cos\frac{t}{2}\right|$$

for any $p(t)\in\mathcal{T}_n$ with equality if and only if

$$p(t) = c\,\frac{\cos\dfrac{2n+1}{2}t}{\cos\dfrac{t}{2}}.$$

We now improve Theorem 6 by considering Lorentz polynomials. Our result is slightly more general, since we consider weights of the form $\left(\frac{\cos t-\cos\omega}{2}\right)^\alpha$ $(0<\alpha\leqq 1)$ instead of $\sqrt{\frac{\cos t-\cos\omega}{2}}$.

THEOREM 7. *Let* $0<\alpha\leqq 1$. *Then* *

$$B_n(\alpha,\omega)=:\sup_{\substack{\sigma(p)\leqq n\\ p\not\equiv 0}}\frac{\|p(t)\|_{I_\omega}}{\left\|p(t)\left(\frac{\cos t-\cos\omega}{2}\right)^\alpha\right\|_{I_\omega}}\sim \tag{36}$$

$$\sim\begin{cases}\left[\dfrac{n(\pi-2\omega)+\sqrt{\omega n}}{2}\right]^\alpha & \text{if}\quad 0<\omega\leqq\pi/2,\\[2ex] \left(\dfrac{1}{2\omega-\pi+\frac{1}{\sqrt{n}}}\right)^\alpha & \text{if}\quad \pi/2\leqq\omega\leqq\pi,\end{cases}$$

and here the supremum is attained if and only if

$$p(t)=c\sin^{2n}\frac{\omega\pm t}{2}.$$

PROOF. Let $0<\tau_n<\min(\omega,\pi-\omega)$ be defined by

$$\sin\tau_n=\frac{n}{n+\alpha}\sin\omega. \tag{37}$$

In estimating $\|p(t)\|_{I_\omega}$, we distinguish two cases.

Case 1. $|\theta|\leqq\tau_n$. Then

$$\frac{|p(\theta)|}{\left\|p(t)\left(\frac{\cos t-\cos\omega}{2}\right)^\alpha\right\|_{I_\omega}}\leqq\left(\frac{2}{\cos t-\cos\omega}\right)^\alpha\leqq \tag{38}$$

$$\leqq\frac{1}{\left(\sin\frac{\omega-\tau_n}{2}\sin\frac{\omega+\tau_n}{2}\right)^\alpha}\qquad(|t|\leqq\tau_n).$$

Case 2. $\tau_n<|\theta|\leqq\omega$, say $\tau_n<\theta\leqq\omega$. An easy calculation shows that the polynomials $\sin^k\frac{\omega-t}{2}\sin^{2n-k}\frac{\omega+t}{2}$ attain their maxima in I_ω at the points $|t_k|\leqq\leqq\min(\omega,\pi-\omega)$ defined by

$$\sin t_k=\frac{n-k}{n}\sin\omega\qquad(0\leqq k\leqq 2n).$$

* $a\sim b$ means that a/b remains between positive bounds depending only on α (but independent of n and ω).

Evidently $t_k < \tau_n$ $(1 \leqq k \leqq 2n)$, whence the polynomials $\sin^k \frac{\omega - t}{2} \sin^{2n-k} \frac{\omega + t}{2}$ $(1 \leqq k \leqq 2n)$ are monotone decreasing in $[\tau_n, \omega]$. Therefore, using the representation (18) (with n instead of m) we get

$$|p(\Theta)| \leqq a_0 \sin^{2n} \frac{\omega+\Theta}{2} + \sum_{k=1}^{2n} a_k \sin^k \frac{\omega-\tau_n}{2} \sin^{2n-k} \frac{\omega+\tau_n}{2} \leqq$$

$$\leqq \left(\frac{\sin \frac{\omega+\Theta}{2}}{\sin \frac{\omega+\tau_n}{2}} \right)^{2n} |p(\tau_n)| \leqq \tag{39}$$

$$\leqq \left(\frac{\sin \frac{\omega+\Theta}{2}}{\sin \frac{\omega+\tau_n}{2}} \right)^{2n} \frac{\left\| p(t) \left(\frac{\cos t - \cos \omega}{2} \right)^{\alpha} \right\|_{I_\omega}}{\left(\sin \frac{\omega-\tau_n}{2} \sin \frac{\omega+\tau_n}{2} \right)^{\alpha}} \qquad (\tau_n \leqq \Theta \leqq \omega),$$

i.e.

$$\frac{|p(\Theta)|}{\left\| p(t) \left(\frac{\cos t - \cos \omega}{2} \right)^{\alpha} \right\|_{I_\omega}} \leqq$$

$$\leqq \left(\frac{\max\limits_{\tau_n \leqq t \leqq \omega} \sin \frac{\omega+t}{2}}{\sin \frac{\omega+\tau_n}{2}} \right)^{2n} \frac{1}{\left(\sin \frac{\omega-\tau_n}{2} \sin \frac{\omega+\tau_n}{2} \right)^{\alpha}} \qquad (\tau_n \leqq \Theta \leqq \omega). \tag{40}$$

Comparing this with (38), we can see that (40) yields a larger upper bound, i.e. the norm is attained in $\tau_n \leqq t \leqq \omega$ (or $-\omega \leqq t \leqq -\tau_n$). It is also seen from (39) that equality holds if and only if $a_1 = a_2 = \ldots = a_{2n} = 0$, i.e. $p(t) = a_0 \sin^{2n} \frac{\omega+t}{2}$. (Considering the interval $[-\omega, -\tau_n]$ we get $p(t) = a_{2n} \sin^{2n} \frac{\omega - t}{2}$ as the extremal polynomial.)

Hence the supremum in (36) is *exactly*

$$B_n(\alpha, \omega) = \left(\frac{\max\limits_{\tau_n \leqq |t| \leqq \omega} \sin \frac{\omega+t}{2}}{\sin \frac{\omega+\tau_n}{2}} \right)^{2n} \cdot \frac{1}{\left(\sin \frac{\omega-\tau_n}{2} \sin \frac{\omega+\tau_n}{2} \right)^{\alpha}}. \tag{41}$$

All we have to do is to calculate the asymptotic value of this expression. (37) can be written in the form

$$2 \sin \frac{\omega-\tau_n}{2} \cos \frac{\omega+\tau_n}{2} = \frac{\alpha \sin \omega}{n+\alpha}, \tag{42}$$

and thus in case $0<\omega\leqq\pi/2$

$$1\leqq\left(\frac{\max\limits_{\tau_n\leqq|t|\leqq\omega}\sin\dfrac{\omega+t}{2}}{\sin\dfrac{\omega+\tau_n}{2}}\right)^{2n}\leqq\left(\frac{\sin\omega}{\sin\dfrac{\omega+\tau_n}{2}}\right)^{2n}=\left(1+\frac{2\sin\dfrac{\omega}{4}\dfrac{\omega-\tau_n}{4}\cos\dfrac{3\omega+\tau_n}{4}}{\sin\dfrac{\omega+\tau_n}{2}}\right)^{2n}\leqq$$

$$\leqq\left(1+\frac{\cos\dfrac{3\omega+\tau_n}{4}\sin\omega}{(n+\alpha)\cos\dfrac{\omega+\tau_n}{2}\sin\dfrac{\omega+\tau_n}{2}}\right)^{2n}\leqq\left(1+\frac{2}{n}\right)^{2n}<e^4,$$

while in case $\pi/2\leqq\omega\leqq\pi$

$$1\leqq\left(\frac{\max\limits_{\tau_n\leqq|t|\leqq\omega}\sin\dfrac{\omega+t}{2}}{\sin\dfrac{\omega+\tau_n}{2}}\right)^{2n}=\left(\sin\frac{\omega+\tau_n}{2}\right)^{-2n}=$$

$$=\left(1+\frac{2\sin^2\dfrac{\pi-\omega-\tau_n}{4}}{\sin\dfrac{\omega+\tau_n}{2}}\right)^{2n}<\left(1+\frac{2\cos^2\dfrac{\omega+\tau_n}{2}}{\sin\dfrac{\omega}{2}}\right)^{2n}<\left(1+\frac{2}{n+\alpha}\right)^{2n}<e^4.$$

Therefore from (41)

$$(43)\quad B_n(\alpha,\omega)\sim\left(\sin\frac{\omega-\tau_n}{2}\sin\frac{\omega+\tau_n}{2}\right)^{-\alpha}\sim\begin{cases}(\omega(\omega-\tau_n))^{-\alpha} & \text{if } 0<\omega\leqq\pi/2,\\ (\omega-\tau_n)^{-\alpha} & \text{if } \pi/2\leqq\omega\leqq\pi.\end{cases}$$

Writing (42) in the form

$$(\pi-2\omega+\omega-\tau_n)(\omega-\tau_n)\sim\frac{\sin\omega}{n},$$

we obtain

$$\omega-\tau_n\sim\sqrt{(2\omega-\pi)^2+\frac{\sin\omega}{n}}+2\omega-\pi\sim$$

$$\sim\begin{cases}\dfrac{\omega}{n(\pi-2\omega)+\sqrt{\omega n}} & \text{if } 0<\omega\leqq\pi/2,\\ 2\omega-\pi+\dfrac{1}{\sqrt{n}} & \text{if } \pi/2\leqq\omega\leqq\pi.\end{cases}$$

This together with (43) yields (36). □

REMARK. We can see from Theorem 7 that for fixed ω

$$B_n(\alpha, \omega) = \begin{cases} o(n^\alpha) & \text{if } 0 < \omega < \pi/2, \\ o(n^{\alpha/2}) & \text{if } \omega = \pi/2, \\ o(1) & \text{if } \pi/2 < \omega \leqq \pi \end{cases}$$

when $n \to \infty$. For $\alpha = \frac{1}{2}$ these estimates are better than those obtained from Theorem 6.

REFERENCES

[1] ERDÉLYI, T. and SZABADOS, J., On polynomials with positive coefficients, *J. Approx. Theory* **54** (1988), 107—122.

[2] LORENTZ, G. G., The degree of approximation by polynomials with positive coefficients, *Math. Ann.* **151** (1963), 239—251. *MR* **27** # 5075.

[3] SCHUR, I., Über das Maximum des absoluten Betrages eines Polynoms in einem gegebenen Intervalle, *Math. Z.* **4** (1919), 271—287.

(Received June 27, 1986)

MTA MATEMATIKAI KUTATÓ INTÉZETE
P.O. BOX 127
H—1364 BUDAPEST
HUNGARY

PRINTED IN HUNGARY
Szegedi Nyomda, Szeged

RECENTLY ACCEPTED PAPERS

MAKAI, E., The full embeddings of the categories of uniform spaces, proximity spaces and related categories into themselves and each other I
PALKA, Z. and RUCIŃSKI, A., Vertex degrees in a random subgraph of a regular graph
MIELOSZYK, E., Boundary value problems for an abstract differential equation
DOMOKOS, G., Digital modelling of chaotic motion
MAKAI, E., The full embeddings of the categories of uniform spaces, proximity spaces and related categories into themselves and each other II
DEÁK, J., On proximity-like relations introduced by F. Riesz in 1908
VÉRTESI, P., On the zeros of Jacobi polynomials
FÉNYES, T., On an operational differential equation
WRESNIK, H., Die Regelflächen des E_n, die eine aus ebenen Kurven bestehende Kongruenzschar tragen I—II
ALPÁR, L., Tauberian theorems for power series of two real variables II
JANKÓ, B., On a unified theory of iteration methods for solving nonlinear operator equations III

CONTENTS

HU ISSN 0081—6906 Index: 26 748

Studia Scientiarum Mathematicarum Hungarica

VOLUME 24
NUMBERS 2—3
1989

1828

AKADÉMIAI KIADÓ, BUDAPEST

STUDIA SCIENTIARUM MATHEMATICARUM HUNGARICA

A QUARTERLY OF THE HUNGARIAN ACADEMY OF SCIENCES

Studia Scientiarum Mathematicarum Hungarica publishes original papers on mathematics mainly in English, but also in German, French and Russian.

Studia Scientiarum Mathematicarum Hungarica is published in yearly volumes of four issues (mostly double numbers published semiannually) by

AKADÉMIAI KIADÓ
Publishing House of the Hungarian Academy of Sciences
H—1117 Budapest, Prielle Kornélia u. 19—35.

Manuscripts and editorial correspondence should be addressed to

J. Merza
Managing Editor
P.O Box 127
H—1364 Budapest

Subscription information

Orders should be addressed to

KULTURA Foreign Trading Company
P.O. Box 149
H—1389 Budapest

or to its representatives abroad.

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 93—117

ALEXANDROVIAN COVER AND SIERPIŃSKIAN EXTENSION

V. K. ZAHAROV

Contents

Introduction

It is well-known which role lower semicontinuous functions play in some sections of mathematics. The family $S_l^*(T)$ of all bounded lower semicontinuous functions on a completely regular space T is the historically first extension of the family $C^*(T)$ of all bounded continuous functions on T. This family is a first stage in a construction of the family of Borel functions and in an extension of measures to the families of measurable functions.

As far as the family $S_l^*(T)$ is not closed under subtraction and multiplication with negative numbers, in 1915 F. Hausdorff ([1] see also [2]—[5]) has introduced

1980 *Mathematics Subject Classification.* Primary 54C10; Secondary 54C30,46E99, 46J10.

Key words and phrases. Alexandrovian and σ-Alexandrovian covers, Sierpińskian and σ-Sierpińskian extensions of $C^*(T)$, perfect preimages, vector lattice and C-ring extensions, Alexandrov determined preimages, almost semicontinuous functions.

the family of bounded *semicontinuous functions* $S^*(T)\equiv\{x-y|x, y\in S_l^*(T)\}$, which is already a lattice algebra. However, in 1921 W. Sierpiński ([6]) has established that the family $S^*(T)$ is not uniformly complete in general and has introduced the uniform completion $\overline{S^*(T)}$ of the family of semicontinuous functions.

The extension $\overline{S^*(T)}$ of the family $C^*(T)$ we shall call the *Sierpińskian extension*. To the very last time neither a functional description of this extension nor its characterization was known.

In this connection the following problems arise: first to select some categories connected with the Sierpińskian extension, secondly to select characteristic properties of this extension within the scope of these categories and third to prove that the selected properties characterize this extension in these categories.

The paper is devoted to a solution of the mentioned problems. Two new categories of extensions of the family $C^*(T)$ are introduced: the category of *vector-lattice extensions inheriting separable decomposition* and the category of *C-ring extensions inheriting separable decomposition.* A series of properties is introduced for objects of the given categories.

For the functional description of the Sierpińskian extension a new class of almost semicontinuous functions is introduced.

In order to prove that the chosen properties characterize the Sierpińskian extension it was required to characterize the *Alexandrovian cover E* of a space T, which can be defined briefly as such a "good" preimage of T that almost semicontinuous functions, lifted on E, become continuous on E. That is why a new topological category of *perfect preimages lifting separable covering* is introduced, some new topological properties of objects of this category are formulated and some characterizations of the Alexandrovian cover are given in the paper (see Theorem 1 and Proposition 3). With the help of these results some vector-lattice and C-ring characterizations of the Sierpińskian extension are given (see Theorems 2 and 3).

Some part of the given paper was published in the paper ([7]). In the paper we shall adhere to the terminology accepted in the books ([8]—[12]).

The author expresses his profound gratitude to Á. Császár and E. Makai Jr. for their help during fulfilling this work at the time of the author's staying in the Budapest University in 1983—84.

§ 1. Basic categories and properties

1.1. Category of perfect preimages lifting separable covering.

We shall suppose that all considered spaces are completely regular and all considered mappings are perfect.

Let T and E be completely regular spaces and ε: $E\to T$ be a surjective perfect mapping.

1.1.1. The preimage E will be called *lower extremally disconnected* if $\mathrm{cl}\,\varepsilon^{-1}G$ is open-closed for any open set G in T.

1.1.2. Let $\mathcal{S}(T)$ denote the set of all non empty countable subsets of the space T. For every countable subset S consider the closed separable subset $T_S\equiv\mathrm{cl}\,S$. The covering $\{T_S|S\in\mathcal{S}(T)\}$ will be called the *separable covering* of the space T.

The preimage E will be called *lifting separable covering* if E has a family of closed subsets $\{E_S|S\in\mathscr{S}(T)\}$ such that $\cup E_S$ is dense in E, $\varepsilon E_S=T_S$ and $S_1\subset S_2$ implies $E_{S_1}\subset E_{S_2}$. The mapping $T_S\mapsto E_S$ will be called the *lifting of the separable covering* and the defined preimage will be denoted by $\{E, \varepsilon\colon E\to T, T_S\mapsto E_S\}$.

A perfect mapping $\gamma\colon E\to\hat{E}$ such that $\varepsilon=\hat{\varepsilon}\circ\gamma$ and $\gamma E_S\subset\hat{E}_S$ will be called a *mapping* (or a *morphism*) *of the preimage* $\{E, \varepsilon\colon E\to T, T_S\mapsto E_S\}$ *into the preimage* $\{\hat{E}, \hat{\varepsilon}\colon \hat{E}\to T, T_S\mapsto\hat{E}_S\}$ and will be denoted by $\{\gamma\}\colon \{E, \varepsilon\colon E\to T, T_S\mapsto E_S\}\to\to\{\hat{E},\hat{\varepsilon}\colon \hat{E}\to T, T_S\mapsto\hat{E}_S\}$.

The preimage $\{E, \varepsilon\colon E\to T, T_S\mapsto E_S\}$ will be called *larger* than the preimage $\{\hat{E}, \hat{\varepsilon}\colon \hat{E}\to T, T_S\mapsto\hat{E}_S\}$ if there exists a mapping $\{\gamma\}\colon \{E, \varepsilon\colon E\to T, T_S\mapsto E_S\}\to\to\{\hat{E}, \hat{\varepsilon}\colon \hat{E}\to T, T_S\mapsto\hat{E}_S\}$ such that the mapping $\gamma\colon E\to\hat{E}$ is surjective and $\gamma E_S=\hat{E}_S$.

1.1.3. Let $\{E, \varepsilon\colon E\to T, T_S\mapsto E_S\}$ be a perfect preimage of T lifting separable covering.

The preimage E will be called *saturated* if for any E_S and any open set G intersecting E_S there exists an E_R such that $\emptyset\neq E_R\subset E_S\cap G$ and $R\subset S$.

The preimage E will be called *filled* if $\cup E_{S_k}$ is dense in E_S for any sequence S_k such that $\cup S_k=S$. Any saturated preimage is filled.

The preimage E will be called *lower disjoined* if $\varepsilon^{-1}G\cap E_S=\emptyset$ implies $\operatorname{cl}\varepsilon^{-1}G\cap\cap E_S=\emptyset$ for any open set G in T.

1.2. Category of vector-lattice extensions inheriting separable decomposition.

We shall suppose that all considered vector lattices are Archimedian, have fixed strong units and are uniformly complete with respect to their units and that all considered vector-lattice homomorphisms preserve these units. Also we shall suppose that all considered vector-lattice ideals are uniformly closed.

Let T be a completely regular space and $C^*(T)$ be the vector lattice of all bounded continuous functions on T. Let X be a vector lattice and $u\colon C^*(T)\to X$ be an injective vector lattice homomorphism. We shall say that X is an *extension* of $C^*(T)$ and shall identify $C^*(T)$ with its image in X.

1.2.1. The extension X will be called *lower Dedekind complete* if any subset of $C^*(T)$, which is bounded above, has a supremum in X.

Let Y be an ideal in X. The ideal Y is called a *component* of X if $y_\xi\in Y$, $x\in X$ and $x=\sup y_\xi$ imply $x\in Y$. The ideal Y will be called a *lower component* of X if $y_\xi\in C^*(T)\cap Y$, $x\in X$ and $x=\sup y_\xi$ imply $x\in Y$.

1.2.2. For any countable set $S\in\mathscr{S}(T)$ consider the ideal $C^*_S(T)\equiv\equiv\{f\in C^*(T)|f(T_S)=0\}$ in $C^*(T)$. The family $\{C^*_S(T)|S\in\mathscr{S}(T)\}$ will be called the *separable decomposition of the vector lattice* $C^*(T)$.

The extension X of $C^*(T)$ will be called *inheriting separable decomposition* if X has a family of proper ideals $\{X_S|S\in\mathscr{S}(T)\}$ such that $\cap X_S=\{0\}$, $uf\in X_S$ iff $f\in C^*_S(T)$ and $S_1\subset S_2$ implies $X_{S_1}\supset X_{S_2}$. The mapping $C^*_S(T)\mapsto X_S$ will be called the *inheritance of separable decomposition* and the defined extension will be denoted by $\{X, u\colon C^*(T)\to X, C^*_S(T)\mapsto X_S\}$.

A vector lattice homomorphism $v\colon X\to\hat{X}$ such that $v\circ u=\hat{u}$ and $vX_S\subset\hat{X}_S$ will be called a *morphism of the extension* $\{X, u\colon C^*(T)\to X, C^*_S(T)\mapsto X_S\}$ *into the extension* $\{\hat{X}, \hat{u}\colon C^*(T)\to\hat{X}, C^*_S(T)\mapsto\hat{X}_S\}$ and will be denoted by $\{v\}\colon \{X, u\colon C^*(T)\to X, C^*_S(T)\mapsto X_S\}\to\{\hat{X}, \hat{u}\colon C^*(T)\to\hat{X}, C^*_S(T)\mapsto\hat{X}_S\}$.

The extension $\{\hat{X}, \hat{u}: C^*(T)\to\hat{X}, C_S^*(T)\mapsto\hat{X}_S\}$ will be called *larger* than the extension $\{X, u: C^*(T)\to X,\ C_S^*(T)\mapsto X_S\}$ if there exists a morphism $\{v\}: \{X, u: C^*(T)\to X, C_S^*(T)\mapsto X_S\}\to\{\hat{X}, \hat{u}: C^*(T)\to\hat{X}, C_S^*(T)\mapsto\hat{X}_S\}$ such that the homomorphism $v: X\to\hat{X}$ is injective and $vx\in\hat{X}_S$ iff $x\in X_S$.

1.2.3. Let $\{X, u: C^*(T)\to X,\ C_S^*(T)\mapsto X_S\}$ be an extension of $C^*(T)$ inheriting separable decomposition.

The extension X will be called *saturated* if for any X_S and any proper component Y such that $Y^d=\{x\in X\mid\forall y\in Y(|x|\cap|y|=0)\}\not\supset X_S$ there exists an X_R such that $X_S\cup Y\subset X_R$ and $R\subset S$.

The extension X will be called *filled* if $\cap X_{S_k}=X_S$ for any sequence S_k such that $\cup S_k=S$.

The extension X will be called *lower component* if every ideal X_S is a lower component of X.

Lemma. *Any saturated extension X is filled.*

Proof. On the strength of Yosida's theorem ([14]) there is a compact E such that the vector lattice X is isomorphic to the vector lattice $C(E)$. Consider the nonempty closed subsets $E_S=\{s\in E\mid\forall x\in X_S(x(s)=0)\}$. Let $S=\cup S_k$. Then $\cup E_{S_k}$ is dense in E_S. In fact assume that there exists an open set G such that $G\cap E_S\neq\emptyset$, $G\cap(\cup E_{S_k})=\emptyset$. Take a regular closed set $F\subset G$ such that $E_S\cap\operatorname{int}F\neq\emptyset$. Consider the proper component $Y\equiv\{y\in X\mid y(F)=0\}$. Then there exists an $R\subset S$ such that $X_S\cup Y\subset X_R$. So $E_R\subset G$. As $R_k\equiv R\cap S_k\neq\emptyset$ for some k we get $\emptyset\neq E_{R_k}\subset E_R\cap E_{S_k}=\emptyset$. From this contradiction we conclude that such set G does not exist.

Now take a $0<x\in\cap X_{S_k}$. Then $x(E_S)=0$. Consider the functions $x_k\equiv\equiv\left(x-\frac{1}{k}\mathbf{1}\right)\vee 0$. From the property $E_S\cap\operatorname{cl}\operatorname{coz}x_k=\emptyset$ we conclude that $x_k\in X_S$. As this ideal is uniformly closed we get $x\in X_S$. The lemma is proved.

1.3. Category of C-ring extensions inheriting separable decomposition.

We shall regard that all considered rings are commutative with a unit and all ring homomorphisms are unitary.

1.3.1. A ring X will be called a *C-ring* if X has the following properties ([15]):

a) for any x, y there exists z such that $x^2+y^2=z^2$;

b) for any x there exist y and z such that $x=y^2-z^2$ and $yz=0$;

c) if for x and for any $n\geqq 1$ there exists $y=y(n)$ such that $n(x^2+y^2)=\mathbf{1}$ then $x=0$;

d) for any x there exists $(\mathbf{1}+x^2)^{-1}$;

e) for any x there exist y and $n\in\mathbf{N}$ such that $x^2+y^2=n\mathbf{1}$;

f) if $\{x_n\}$ is a sequence such that for any $k\geqq 1$ there exists $n_0=n_0(k)$ such that $m, n\geqq n_0$ implies $\exists y,\ k((x_m-x_n)^2+y^2)=\mathbf{1}$ then there exists x such that for any $k\geqq 1$ there exists $m=m(k)$ and z such that $k((x-x_m)^2+z^2)=\mathbf{1}$.

A ring ideal Y of the C-ring X will be called *C-ideal* if for any sequence $\{y_n\}$ and any x such that for any $k\geqq 1$ there exists $m=m(k)$ and z such that $k((x-y_m)^2+z^2)=\mathbf{1}$, the condition $\{y_n\}\subset Y$ implies $x\in Y$.

The importance of the class of C-rings follows from the following

THEOREM (Delfosse) ([15]). *A commutative ring X with a unit is a C-ring iff X is isomorphic to a ring $C(K)$ of all continuous functions on some compact space K.*

COROLLARY. *With respect to an order defined by the cone $P \equiv \{x \in X \mid \exists y \in X(x = y^2)\}$ the C-ring X is a lattice ring and the ring isomorphism $X \approx C(K)$ is a lattice ring isomorphism.*

1.3.2. Let T be a completely regular space and $C^*(T)$ be the C-ring of all bounded continuous functions on T. Let X be a C-ring and $u\colon C^*(T) \to X$ be an injective ring homomorphism. We shall say that X is a *C-ring extension* of $C^*(T)$ and shall identify $C^*(T)$ with its image in X. If Y and Z are modules over the C-ring X then the set of all module homomorphisms from Y into Z is denoted by $\mathrm{Hom}_X(Y, Z)$. Let Y and Z be ring ideals in the C-ring X. A homomorphism $g \in \mathrm{Hom}_X(Y, Z)$ will be called bounded if there is a natural number n such that $|gy| \leqq n|y|$ for any $y \in Y$. The subset of $\mathrm{Hom}_X(Y, Z)$ consisting of all the bounded homomorphisms will be denoted by $\mathrm{Hom}^*_X(Y, Z)$.

The first and second annihilator of a subset Y of X will be denoted as usual by Y^* and Y^{**}, resp.

The extension X will be called *lower continuing* if for any ring ideal Y of the ring $C^*(T)$ and for any homomorphism $g \in \mathrm{Hom}^*_{C^*(T)}(Y, C^*(T) \cap Y^{**})$ there exists a homomorphism $h \in \mathrm{Hom}^*_X(X, Y^{**})$ extending g.

A ring ideal Z in X will be called a *lower segment* of X if for any ring ideal Y of the ring $C^*(T)$ and for any pair of homomorphisms $g \in \mathrm{Hom}^*_{C^*(T)}(Y, C^*(T) \cap Y^{**})$ and $h \in \mathrm{Hom}^*_X(X, Y^{**})$ such that h extends g the condition $gY \subset Z$ implies $hX \subset Z$.

1.3.3. For any countable set $S \in \mathscr{S}(T)$ consider the C-ideal $C^*_S(T) \equiv \{f \in C^*(T) \mid f(T_S) = 0\}$ in the C-ring $C^*(T)$. The family $\{C^*_S(T) \mid S \in \mathscr{S}(T)\}$ will be called the *separable decomposition* of the C-ring $C^*(T)$.

The extension X of $C^*(T)$ will be called *inheriting separable decomposition* if X has a family of proper C-ideals $\{X_S \mid S \in \mathscr{S}(T)\}$ such that $\bigcap X_S = \{0\}$, $uf \in X_S$ iff $f \in C^*_S(T)$ and $S_1 \subset S_2$ implies $X_{S_1} \supset X_{S_2}$. The mapping $C^*_S(T) \mapsto X_S$ will be called the *inheritance of separable decomposition* and the defined extension will be denoted by $\{X, u\colon C^*(T) \to X,\ C^*_S(T) \mapsto X_S\}$.

A ring homomorphism $v\colon X \to \hat{X}$ such that $v \circ u = \hat{u}$ and $vX_S \subset \hat{X}_S$ will be called a *morphism of the extension* $\{X, u\colon C^*(T) \to X,\ C^*_S(T) \to X_S\}$ *into the extension* $\{\hat{X}, \hat{u}\colon C^*(T) \to \hat{X}, C^*_S(T) \mapsto \hat{X}_S\}$ and will be denoted by $\{v\}\colon \{X, u\colon C^*(T) \to X, C^*_S(T) \mapsto X_S\} \to \{\hat{X}, \hat{u}\colon C^*(T) \to \hat{X}, C^*_S(T) \mapsto \hat{X}_S\}$.

The partial preorder on C-ring extensions of $C^*(T)$ inheriting separable decomposition is defined as in Section 1.2.2.

1.3.4. Let $\{X, u\colon C^*(T) \to X, C^*_S(T) \mapsto X_S\}$ be an extension of $C^*(T)$ inheriting separable decomposition.

A ring ideal Y in X is called an *annihilator ideal* if Y coincides with its own second annihilator Y^{**}.

The extension X will be called *saturated* if for any X_S and any proper annihilator ring ideal Y such that $Y^* \not\subset X_S$ there exists an X_R such that $X_S \cup Y \subset X_R$ and $R \subset S$.

The extension X will be called *filled* if $\bigcap X_{S_k} = X_S$ for any sequence S_k such that $\bigcup S_k = S$. Any saturated extension is filled.

The extension X will be called *lower segment* if any X_S is a lower segment of X.

§ 2. Alexandrovian and σ-Alexandrovian covers

Let T be a completely regular space. Let $\mathscr{A}(T)$ denote the field of subsets of T, generated by all open sets. This field will be called the *field of Alexandrov subsets* of T[1].

2.1. Construction and properties of Alexandrovian cover

Consider the Stone compact E_0 of all ultrafilters in $\mathscr{A}(T)$. For any point $s \in E_0$ let P_s denote the set $\bigcap \{\mathrm{cl}\, A | A \in \Theta_s\}$ where s corresponds to the ultrafilter Θ_s. Consider the subspace $E \equiv \{s \in E_0 | P_s \neq \emptyset\}$ and define the surjective continuous mapping $\varepsilon: E \to T$ such that $\varepsilon s \equiv P_s$. The space E with the mapping ε will be called the Alexandrovian cover of T.

Let i_0 be the Stone isomorphism between $\mathscr{A}(T)$ and the Boolean algebra $\Delta(E_0)$ of all open-closed subsets of E_0. Let $i: \mathscr{A}(T) \to \Delta(E)$ be the corresponding injective homomorphism of Boolean algebras such that $iA \equiv E \cap i_0 A$.

Lemma 1. *For any set $A \in \mathscr{A}(T)$ there are open sets G_k and closed sets F_k such that $A = \bigcup_k (G_k \cap F_k)$.*

Proof. It is clear that the family of all sets of such form is closed under joins and contains the set T. Let $A = \bigcup \{G'_k \cap F'_k | k = 1, \dots, n\}$. Consider the sets $F_k \equiv T \setminus G'_k$ and $G_k \equiv T \setminus F'_k$. Then $T \setminus A = \bigcap \{G_k \cup F_k\}$. It is clear that $T \setminus A \subset \subset \bigcup \{(G_{i_1} \cap \dots \cap G_{i_l}) \cap (F_{j_1} \cap \dots \cap F_{j_m}) | \{i_1, \dots, i_l\} \cup \{j_1, \dots, j_m\} = \{1, \dots, n\}\}$. Conversely, consider the set $C \equiv (G_{i_1} \cap \dots \cap G_{i_l}) \cap (F_{j_1} \cap \dots \cap F_{j_m})$. Take an arbitrary k. If $k = i_p$ then $C \subset G_{i_p} \subset G_k \cup F_k$. If $k = j_q$ then $C \subset F_{j_q} \subset G_k \cup F_k$. Therefore $C \subset T \setminus A$. Thus $T \setminus A$ is the set of mentioned form. Consequently, such a family is a field. The lemma is proved.

Lemma 2. *The subspace E is dense in E_0, the homomorphism i is injective and the mapping $\varepsilon: E \to T$ is perfect.*

Proof. Consider an arbitrary open-closed set $U \equiv i_0 A$ in E_0. For some point $t \in A$ consider the ultrafilter $\Theta \equiv \{A' \in \mathscr{A}(T) | t \in A'\}$. Let a point $s \in E_0$ correspond to Θ. Then $s \in E \cap U$. Consequently, E is dense in E_0. Since $iA = E \cap i_0 A$ the homomorphism i is injective.

Denote βT by S. For any point $s \in E_0$ consider the set $P'_s \equiv \bigcap \{\mathrm{cl}_S\, A | A \in \Theta_s\} \neq \emptyset$ where s corresponds to the ultrafilter Θ_s. Assume that t_1 and t_2 are different points from P'_s. Then there exists an open set G in S such that $t_1 \in G$ and $t_2 \notin \mathrm{cl}_S\, G$. Let $G' \equiv G \cap T$. Since $G' \cap A \neq \emptyset$ for any $A \in \Theta_s$ we have $G' \in \Theta_s$. Therefore $t_2 \in \mathrm{cl}_S\, G'$ but this is false. Hence P'_s consists of only one point. Define the mapping $\varepsilon_0: E_0 \to S$ setting $\varepsilon_0 s \equiv P'_s$. This mapping is continuous. In fact, let U be a neighbourhood of the point $\varepsilon_0 s$ in S. Take an open neighbourhood G of this point such that $\mathrm{cl}\, G \subset U$. Consider the sets $G' \equiv G \cap T$ and $V \equiv i_0 G'$. Since $G' \cap A \neq \emptyset$ for any $A \in \Theta_s$ we have $G' \in \Theta_s$ and therefore $s \in V$. Besides $\varepsilon_0 V \subset U$.

[1] This term is introduced in honour of A. D. Alexandrov, who has used the field $\mathscr{A}(T)$ for extending Riesz' theorem from compact to normal spaces ([16]).

It is clear that $\varepsilon \equiv \varepsilon_0 | E$ and $E \subset \varepsilon_0^{-1} T$. Let $\varepsilon_0 s \in T$. Then $P_s = \bigcap \{T \cap \mathrm{cl}_S A | A \in \Theta_s\} = T \cap \varepsilon_0 s \neq \emptyset$. Hence $s \in E$. Thus $E = \varepsilon_0^{-1} T$. This implies that ε is perfect. The lemma is proved.

Associate with a countable set S the closed subspace E_S of all the ultrafilters from E, all the members of which intersect the set S.

Alexandrov sets A_1 and A_2 will be called *S-equivalent* if $(A \triangle A') \cap S = \emptyset$. Let $\mathscr{A}_S(T)$ denote the Boolean algebra of all classes of S-equivalence $\bar{A}$ of elements A from $\mathscr{A}(T)$. Let $(A \triangle A') \cap S = \emptyset$. Let $s \in iA \cap E_S$ and s corresponds to an ultrafilter Θ_s. Assume $A' \notin \Theta_s$. Then $A \triangle A' \supset A \setminus A' \in \Theta_s$ but this is false. Hence $s \in iA' \cap E_S$. Thus $iA \cap E_S = iA' \cap E_S$. So we can define correctly the homomorphism of Boolean algebras $i_S \colon \mathscr{A}_S(T) \to \triangle(E_S)$ such that $i_S \bar{A} \equiv iA \cap E_S$.

LEMMA 3. *The family* $\Delta_0(E_S) \equiv \{i_S \bar{A} | A \in \mathscr{A}(T)\}$ *is a base in the space* E_S, *the homomorphism* $i_S \colon \mathscr{A}_S(T) \to \Delta_0(E_S)$ *is injective and* $\varepsilon E_S = T_S$.

PROOF. Denote by E_{0S} the subspace of E_0 consisting of all the ultrafilters, all the members of which intersect the set S. Consider the factor-homomorphism $h_S \colon \mathscr{A}(T) \to \mathscr{A}_S(T)$. Let $i_S \bar{A} = \emptyset$ and assume $A \cap S \neq \emptyset$. Take a compact set $F \subset A \cap S$. Consider the proper filter base Θ_0 in $\mathscr{A}(T)$ consisting of the set A and all open sets G containing F. Then $\Theta_0' \equiv h_S \Theta_0$ is a proper filter base in $\mathscr{A}_S(T)$. Imbed Θ_0' in some ultrafilter Θ' and consider the ultrafilter $\Theta \equiv h_S^{-1} \Theta'$. Let s be the point in E_0 corresponding to Θ. Then $s \in E_{0S}$. If $A' \in \Theta$ then $F \cap \mathrm{cl}\, A' \neq \emptyset$. Otherwise $G \cap \mathrm{cl}\, A' = \emptyset$ for some $G \in \Theta_0 \subset \Theta$. This implies $A' \cap G \in \Theta$. Hence $\emptyset = A' \cap G \cap S \neq \emptyset$. It follows from this contradiction that $\bigcap \{F \cap \mathrm{cl}\, A' | A' \in \Theta\} \neq \emptyset$. Therefore $s \in E_S$. Besides $A \in \Theta$ implies $s \in i_S A = \emptyset$. From this contradiction we conclude that $\bar{A} = 0$.

Let $t \in T_S$. Consider the proper filter base Θ_0 consisting of all open sets G containing t. Imbed $h_S \Theta_0$ in some ultrafilter Θ' and consider the ultrafilter $\Theta \equiv h_S^{-1} \Theta'$. Let s be the corresponding point. As $t \in \bigcap \{\mathrm{cl}\, A | A \in \Theta\}$ we have $s \in E_S$ and $\varepsilon s = t$. Further let a point $s \in E_S$ correspond to an ultrafilter Θ_s and assume that $t \equiv \varepsilon s \notin T_S$. Take a cozero set C such that $t \in C \subset T \setminus T_S$. As $C \cap A \neq \emptyset$ for any $A \in \Theta$ we obtain $C \in \Theta$. So $C \cap S \neq \emptyset$ but this is false. This contradiction means that $t \in T_S$. The lemma is proved.

LEMMA 4. *For any* E_S *and any open-closed set* $\emptyset \neq U \equiv iA$ *in* E *there exists a set* $R \subset S$ *such that* $E_R \cap U \subset E_S$.

PROOF. Consider the non-empty set $V \equiv i_S \bar{A}$. Then the set $R \equiv A \cap S$ is non-empty. We have $E_R \subset E_S$. Let s be a point from V corresponding to an ultrafilter Θ. Then $A \in \Theta$. Therefore for any $A' \in \Theta$ we have $A' \cap R = A \cap A \cap S \neq \emptyset$. This means $s \in E_R$. Conversely, let $s \in E_R$. For any $A' \in \Theta$ we have $A' \cap A \cap S \neq \emptyset$. It has as a consequence that $A \in \Theta$ and so $s \in V$. The lemma is proved.

LEMMA 5. *Let* A *be an Alexandrov set. Then* $iA \triangle \varepsilon^{-1} A \not\supset E_S$ *for any* S.

PROOF. Assume that $E_S \subset iA \triangle \varepsilon^{-1} A$ for some S. Suppose that $E_S \cap iA \neq \emptyset$. According to the previous lemma there exists a set $R \subset S$ such that $E_R \subset iA \setminus \varepsilon^{-1} A$. This implies $T_R \cap A = \emptyset$. Consequently, $iA \cap E_R = \emptyset$. But this is false. It follows from this contradiction that $E_S \subset \varepsilon^{-1} A \setminus iA$. Then $T_S \subset A$ implies by Lemma 3 $iA \cap E_S \neq \emptyset$ but this is also false. The lemma is proved.

LEMMA 6. *If* $S_1\cap S_2=\emptyset$ *then* $E_{S_1}\cap E_{S_2}$ *is a nowhere dense set in the subspaces* E_{S_1} *and* E_{S_2}.

PROOF. For any point $s_k\in S_1$ take a cozero set C_k such that $S_2\subset C_k$ and $s_k\notin C_k$. Then $\bigcap C_k\cap S_1=\emptyset$. Let s be a point in $E_{S_1}\cap E_{S_2}\equiv F$ corresponding to an ultrafilter Θ. As $T\setminus C_k\notin\Theta$ for any k we get $C_k\in\Theta$ and hence $s\in\bigcap iC_k$. Therefore $F\subset\bigcap i_{S_1}C_k\equiv H$. Assume that the set H is not nowhere dense in E_{S_1}. By Lemma 3 $\emptyset\neq i_{S_1}\bar{A}\subset H$ for some set A. Consequently, $0<\bar{A}\subset\bar{C}_k$ in $\mathscr{A}_{S_1}(T)$ for any k. Therefore $\emptyset\neq A\cap S_1\subset\bigcap C_k\cap S_1=\emptyset$ but this is false. Thus H and F are nowhere dense. The lemma is proved.

LEMMA 7. *For any* S_1 *and* S_2 *we have* $E_{S_1}\cup E_{S_2}=E_{S_1\cup S_2}$.

PROOF. Denote $S_1\cup S_2$ by S and assume that $G\equiv E_S\setminus(E_{S_1}\cup E_{S_2})$ is non-empty. By Lemma 3 $\emptyset\neq i_S\bar{A}\subset G$ for some set A. By Lemma 4 $i_S\bar{A}=E_R$ for some set $R\subset S$. If $R\cap S_1\neq\emptyset$ then $E_R\cap E_{S_1}\neq\emptyset$ but this is false. Then $R\cap S_2\neq\emptyset$ implies $E_R\cap E_{S_2}\neq\emptyset$ but this is also false. From this contradiction we conclude that $G=\emptyset$. The lemma is proved.

LEMMA 8. *If* $S_1\subset S_2$ *then* $E_{S_1}\subset E_{S_2}$ *and the set* E_{S_1} *has a non-empty interior in the space* E_{S_2}.

PROOF. Consider the set $S\equiv S_2\setminus S_1$. By the previous lemma $E_{S_2}=E_S\cup E_{S_1}$. Assume that E_{S_1} has an empty interior in E_{S_2}. As $E_{S_2}\setminus E_S$ is contained in this interior we get $E_{S_2}=E_S$. According to Lemma 6 the set $E_{S_1}=E_{S_1}\cap E_S$ is nowhere dense in the space E_{S_1} but this is impossible. The proof is finished.

Consider on E the base $\mathscr{C}(E)$ of all cozero sets.

LEMMA 9. *For any open subset* G *of* T *we have* $\mathrm{cl}\,\varepsilon^{-1}G=iG$.

PROOF. Assume that there exists a point $s\in\varepsilon^{-1}G\cap i(T\setminus G)$. Then $\varepsilon s\in T\setminus G$ but this is impossible. On the other hand assume that there exists a non-empty set $U\equiv iA$ such that $U\subset iG\setminus\mathrm{cl}\,\varepsilon^{-1}G$. We can suppose that $A\subset G$. Let $A=\bigcup(G_k\cap F_k)$. Then $G_k\cap F_k\neq\emptyset$ for some index. Therefore there exists a closed set $F\subset G_k\cap F_k$. Consider the set $V\equiv iF\neq\emptyset$. Then $\emptyset\neq\varepsilon V\subset F\subset G$. But this contradicts to the inclusion $V\subset\varepsilon^{-1}(T\setminus G)$.

COROLLARY 1. *The preimage* E *is lower extremally disconnected.*

COROLLARY 2. *The preimage* E *is lower disjoined.*

PROOF. Let $\varepsilon^{-1}G\cap E_S=\emptyset$. Then $G\cap S=\emptyset$ implies $iG\cap E_S=\emptyset$.

2.2. Alexandrov determined preimages.

Now let E be a saturated preimage of T lifting separable covering.

Let P and Q be subsets of E. If $P\setminus Q\not\supset E_S$ for any S we shall say that P is *almost contained in* Q and write $P\underset{\sim}{\subset}Q$.

Let $\{C_k\}\subset\mathscr{C}(E)$ be a finite covering of the space E and $\{A_k\}\subset\mathscr{A}(T)$ be a finite covering of the space T. The family $\{C_k,\varepsilon^{-1}A_k\}$ will be called a *cohesive covering* of the space E if $\varepsilon^{-1}A_k\underset{\sim}{\subset}C_k$ for any k. It is clear that the covering $\{A_k\}$ can be supposed to be disjoint.

LEMMA 10. *Let* $\{C_j, \varepsilon^{-1}A_j\}$ *and* $\{C_k, \varepsilon^{-1}A_k\}$ *be cohesive coverings of* E. *Then the family* $\{C_j \cap C_k,\ \varepsilon^{-1}(A_j \cap A_k)\}$ *is a cohesive covering, too.*

PROOF. Assume that $\varepsilon^{-1}(A_j \cap A_k) \setminus (C_j \cap C_k) \supset E_S$ for some S. If $E_S \cap C_j = \emptyset$ then $\varepsilon^{-1}A_j \setminus C_j \supset E_S$ but this is impossible. Consequently, $E_S \cap C_j \supset E_R$ for some R. Then $E_R \subset \varepsilon^{-1}A_k \setminus C_k$ but this is impossible, too. It follows from this contradiction that such E_S does not exist. The lemma is proved.

An entourage $\cup (C_k \times C_k)$ of the diagonal in the space $E \times E$ generated by a cohesive covering $\{C_k, \varepsilon^{-1}A_k\}$ will be called a *c-entourage*. The family of all c-entourages will be denoted by $\mathscr{V}_c$.

Let $\mathscr{U}$ be some family of entourages in $E \times E$. We shall say that $\mathscr{V}_c$ is finer than $\mathscr{U}$ if for any entourage $U \in \mathscr{U}$ there exists a c-entourage $V \in \mathscr{V}_c$ such that $V \subset U$.

The preimage E will be called *Alexandrov determined* if there is a uniformity $\mathscr{U}$ on E generating the original topology and such that $\mathscr{V}_c$ is finer than $\mathscr{U}$.

Let E be the Alexandrovian cover of T. Consider on $E \times E$ the family $\mathscr{U}$ of all entourages U such that for U there exists a cohesive covering $\{C_k, \varepsilon^{-1}A_k\}$ such that $\cup (\mathrm{cl}\, C_k \times \mathrm{cl}\, C_k) \subset U$.

LEMMA 11. *Let* E *be the Alexandrovian cover of* T. *Then the preimage* E *is Alexandrov determined.*

PROOF. It follows from the previous lemma that $\mathscr{U}$ is a filter. It is clear that $U^{-1} \in \mathscr{U}$ for any $U \in \mathscr{U}$. Let for an entourage U there exist a corresponding cohesive covering $\{C_k, \varepsilon^{-1}A_k\}$ with a partition $\{A_k\}$. Consider the new covering $\{U_k \equiv iA_k\}$ of E. By virtue of Lemma 5 the family $\{U_k, \varepsilon^{-1}A_k\}$ is a cohesive covering. Consider a c-entourage $V \equiv \cup (U_k \times U_k)$. Then $V \in \mathscr{U}$.

Assume that $P \equiv U_k \setminus \mathrm{cl}\, C_k \neq \emptyset$. Then there exists an S such that $E_S \subset P$. Therefore $A_k \cap S \neq \emptyset$ and consequently there exists a compact set $R \subset A_k \cap S$. This implies $E_R \subset \varepsilon^{-1}A \cap E_S \subset \varepsilon^{-1}A_k \setminus C_k$ but this is impossible. Hence $U_k \subset \mathrm{cl}\, C_k$. Thus $V \subset U$. But it follows from the disjointness of squares of the entourage V that $V \circ V = V$. Therefore $V \circ V \subset U$. So $\mathscr{U}$ is a uniformity. It is clear that $\mathscr{V}_c$ is finer than $\mathscr{U}$.

Let s be an arbitrary point in E and G be an arbitrary neighbourhood of this point in the original topology T_0. Then $s \in U_1 \subset G$ for some set $U_1 \equiv iA$. Let $U_2 \equiv i(T \setminus A)$. Consider the c-entourage $U \equiv \cup (U_k \times U_k) \in \mathscr{U}$. The set

$$H \equiv \{r \in E \mid (s, r) \in U\}$$

is a neighbourhood of s in the topology T_1 generated by the uniformity $\mathscr{U}$. As $H \subset G$ the topology T_1 is finer than the topology T_0. Vice versa let H be an arbitrary neighbourhood of s in T_1. Then $H \equiv \{r \in E \mid (s, r) \in U\}$ for some entourage $U \in \mathscr{U}$. For U there exists a cohesive covering $\{C_k, \varepsilon^{-1}A_k\}$ such that $\cup \{\mathrm{cl}\, C_k \times \mathrm{cl}\, C_k) \subset U$. Then $s \in C_k \subset H$ for some k. Consequently, T_0 is finer than T_1. Thus these topologies coincide. The lemma is proved.

2.3. Functional description of Alexandrov determined preimages.

Now we shall give a functional description of Alexandrov determined preimages. Let E be a perfect saturated preimage of T lifting separable covering.

Denote by $C^*(E, T, \varepsilon)$ the set of all bounded continuous functions f on E such that for any n there exists a cohesive covering $\{C_k, \varepsilon^{-1}A_k\}$ such that the oscillation $\omega(f, C_k)$ of the function f on any set C_k is less than $u_n \equiv 1/n$.

LEMMA 12. *Let $\{C_k, \varepsilon^{-1}A_k\}$ be a cohesive covering of E. Then the set $\bigcup(C_k \cap \varepsilon^{-1}A_k)$ is dense in E.*

PROOF. Denote this set by H. Assume that there exists an open set G such that $G \cap H = \emptyset$. Then by virtue of the saturatedness of E there exists some $E_S \subset G$. As $A_k \cap S \neq \emptyset$ for some K there exists some compact set $K \subset A_k \cap S$. We have $E_K \subset \subset \varepsilon^{-1}A_k \cap E_S \subset \varepsilon^{-1}A_k \setminus C_k$ but this is impossible. Thus $\mathrm{cl}\, H = E$. The lemma is proved.

LEMMA 13. *Let $f \in C^*(E)$ and $\{C_k, \varepsilon^{-1}A_k\}$ be a cohesive covering such that $\omega(f, C_k \cap \varepsilon^{-1}A_k) < u_n$. Then there exists a cohesive covering $\{D_k, \varepsilon^{-1}A_k\}$ such that $\omega(f, D_k) < 3u_n$.*

PROOF. Denote $R_k \equiv C_k \cap \varepsilon^{-1}A_k$ and $F_k \equiv \mathrm{cl}\, R_k$. Then $\bigcup F_k = E$. Divide an interval containing the range of the function f by points a_j so that $a_{j+1} - a_j < u_n/3$. Consider the covering $G_j \equiv f^{-1}(]a_{j-1}, a_{j+2}])$ of E. Let $J_k \equiv \{j \mid G_j \cap F_k \neq \emptyset\}$. Consider the cozero sets $D_k \equiv \bigcup\{G_j \mid j \in J_k\}$. It is clear that $F_k \subset D_k$. Therefore $\{D_k\}$ is a covering of E. Let $r, s \in D_k$. Then $r \in G_i$ and $s \in G_j$ for some $i, j \in J_k$. For i and j there exist $r_1 \in G_i \cap R_k$ and $s_1 \in G_j \cap R_k$. Therefore we have $|f(r) - f(s)| \leq \leq |f(r) - f(r_1)| + |f(r_1) - f(s_1)| + |f(s_1) - f(s)| < 3u_n$. Consequently, $\omega(f, D_k) < 3u_n$. Since $\varepsilon^{-1}A_k \setminus D_k \subset \varepsilon^{-1}A_k \setminus C_k$ the covering $\{D_k, \varepsilon^{-1}A_k\}$ is cohesive. The lemma is proved.

LEMMA 14. *The family $C^*(E, T, \varepsilon)$ is a uniformly complete lattice algebra.*

PROOF. Denote $C^*(E, T, \varepsilon)$ by Φ and the set of all step-functions $y \equiv \sum\{a_k \chi(A_k) \mid A_k \in \mathscr{A}(T),\ A_k \cap A_j = \emptyset\}$ by Y.

Let $f \in \Phi$. Then for every n there exists a cohesive covering $\{C_k, \varepsilon^{-1}A_k\}$ such that $\omega(f, C_k) < u_n$. Consider the numbers $a_k \equiv \inf\{f(s) \mid s \in C_k\}$ and the step-function $y \equiv \sum a_k \chi(A_k) \in Y$. Let $P_n \equiv \bigcup(\varepsilon^{-1}A_k \setminus C_k)$. Then $|f(s) - y_n \circ \varepsilon(s)| < u_n$ for any $s \notin P_n$.

Conversely, let for an $f \in C^*(E)$ and any n there exist a cohesive covering $\{C_k, \varepsilon^{-1}A_k\}$ and a step-function $y_n \equiv \sum a_k \chi(A_k) \in Y$ such that $|f(s) - y_n \circ \varepsilon(s)| < u_n$ for any $s \notin P_n \equiv \bigcup(\varepsilon^{-1}A_k \setminus C_k)$. Let $r, s \in \varepsilon^{-1}A_k \cap C_k$. Then it follows from the disjointness of $\{A_k\}$ that $r, s \notin P_n$. Therefore $|f(r) - f(s)| < 2u_n$. By virtue of the previous lemma we have $f \in \Phi$.

Let $f \in C^*(E)$ and $f_n \in \Phi$ be a sequence such that $|f(s) - f_n(s)| < v_n$ for all $s \in E$ and for some sequence of real numbers v_n decreasing to zero. Take a number m such that $v_m < u_n/2$. For f_m consider a cohesive covering $\{C_k, \varepsilon^{-1}A_k\}$ and a step-function $y_n \equiv \sum a_k \chi(A_k)$ such that $|f_m(s) - y_n \circ \varepsilon(s)| < u_n/2$ for any $s \notin P_n$. Then $|f(s) - y_n \circ \varepsilon(s)| < u_n$ implies $f \in \Phi$. This means the uniform completeness.

Let $f, g \in \Phi$. Then for them there exist cohesive coverings $\{C_k, \varepsilon^{-1}A_k\}$ and $\{C_j, \varepsilon^{-1}A_j\}$ and step-functions $y_n \equiv \sum a_k \chi s(A_k)$ and $z_n \equiv \sum a_j \chi(A_j)$ such that $|f(s) - - y_n \circ \varepsilon(s)| < u_n/2$ for any $s \notin P_n$ and $|g(s) - z_n \circ \varepsilon(s)| < u_n/2$ for any $s \notin Q_n \equiv \bigcup(\varepsilon^{-1}A_j \setminus \setminus C_j)$. Consider the cohesive covering $\{C_k \cap C_j, \varepsilon^{-1}(A_k \cap A_j)\}$ and the step-function $y_n \vee z_n = \sum(a_k \vee a_j)\chi(A_k \cap A_j)$. Let $s \notin R_n \equiv \bigcup(\varepsilon^{-1}(A_k \cap A_j) \setminus (C_k \cap C_j))$. Then

$s\in\varepsilon^{-1}(A_k\cap A_j)\cap(C_k\cap C_j)$ for some indexes. Consequently, $s\in\varepsilon^{-1}A_k\cap C_k$ implies $s\notin P_n$. Similarly, $s\notin Q_n$. This gives $|f(s)\vee g(s)-(y_n\vee z_n)\circ\varepsilon(s)|\leqq|f(s)\vee g(s)-$ $-y_n\circ\varepsilon(s)\vee g(s)|+|y_n\circ\varepsilon(s)\vee g(s)-(y_n\vee z_n)\circ\varepsilon(s)|<u_n$. Hence $f\vee g\in\Phi$. In the same way the presence of the addition is established. The lemma is proved.

The above mentioned properties lead us to the following

PROPOSITION 1. *Let E be a perfect saturated preimage of T lifting separable covering. Then the following assertions are equivalent:*

a) *E is Alexandrov determined;*

b) *the family $\mathscr{C}_0(E)\equiv\{\operatorname{coz}f\,|\,f\in C^*(E,T,\varepsilon)\}$ is a base of the topology on E.*

PROOF. a)$\Rightarrow$b). Denote $C^*(E,T,\varepsilon)$ by Φ. Consider on E the uniformly complete vector lattice Φ_0 of all bounded uniformly continuous functions with respect to the uniformity $\mathscr{U}$ on E. Let s be an arbitrary point in E and G be an arbitrary open neighbourhood of this point in the original topology T_0. As the topology T_1, generated by the uniformity $\mathscr{U}$, is finer than the topology T_0 there exists a T_1-neighbourhood H of s such that $s\in H\cup G$. For H there exists an entourage $V\in\mathscr{U}$ such that $H=\{r\in E\,|\,(s,r)\in V\}$. Let $F\equiv E\setminus G$. Since $(\{s\}\times F)\cap V=\emptyset$ the sets $\{s\}$ and F are remote with respect to the proximity generated by the given uniformity. Therefore there exists some function $f\in\Phi_0$ such that $f(s)=1$ and $f(F)=\{0\}$ (see [15], 4.2.f.2).

Check that $\Phi_0\subset\Phi$. Let $f\in\Phi_0\subset C^*(E)$. Then for every n there exists an entourage $V_n\in\mathscr{U}$ such that $(r,s)\in V_n$ implies $|f(r)-f(s)|<u_n$. For V_n there exists a cohesive covering $\{C_k,\varepsilon^{-1}A_k\}$ with a partition $\{A_k\}$ such that $\bigcup(C_k\times C_k)\subset V_n$. Consider the numbers $a_k\equiv\inf\{f(r)\,|\,r\in C_k\}$ and the step-function $y_n\equiv\sum a_k\chi(A_k)$. Let an $s\notin P_n\equiv\bigcup(\varepsilon^{-1}A_k\setminus C_k)$. Then $s\in\varepsilon^{-1}A_k\cap C_k$ for some k. Therefore $|f(s)-$ $-y_n\circ\varepsilon(s)|\leqq u_n$. Now it follows from the proof of the previous lemma that $f\in\Phi$.

b)$\Rightarrow$a) For every function $f\in\Phi$ define a pseudo-metric p_f by the equality $p_f(r,s)\equiv|f(r)-f(s)|$. Consider the uniformity $\mathscr{U}$ on E defined by the pseudo-metrics $\{p_f\,|\,f\in\Phi\}$. Take an arbitrary basic entourage $U\equiv\{(r,s)\in E\times E\,|\,p_{f_k}(r,s)<$ $<\varepsilon,\ k=1,\ldots,m\}$. Take a number n such that $u_n<\varepsilon$. Then for the function f_k there exists a cohesive covering $\{C^k_{j_k},\varepsilon^{-1}A^k_{j_k}\,|\,j_k=1,\ldots,n_k\}$ such that $\omega(f_k,C^k_{j_k})<u_n$. Consider the new cohesive covering $\{C^1_{j_1}\cap\ldots\cap C^m_{j_m},\ \varepsilon^{-1}(A^1_{j_1}\cap\ldots\cap A^m_{j_m})\,|\,j_k\leqq n_k,$ $k\leqq m\}$ and the entourage V determined by this covering. Let $(r,s)\in V$. Then $p_{f_k}(r,s)<\varepsilon$ for any k. Consequently, $V\subset U$.

Further let s be an arbitrary point in E and G be an arbitrary neighbourhood of this point in the original topology T_0. Then there exists a function $f\in\Phi$ with the cozero set C such that $f(E)\subset[0,1]$, $f(s)=1$ and $C\subset G$. Consider the basic entourage $U\equiv\{(r,t)\in E\times E\,|\,p_f(r,t)<1/2\}$. The set $H\equiv\{r\in E\,|\,(s,r)\in U\}$ is a neighbourhood of s in the topology T_1 generated by the defined uniformity. As $H=$ $=f^{-1}(]1/2,1])\subset C$ then T_1 is finer than T_0.

Conversely, let H be an arbitrary neighbourhood of s in T_1. Then there exists an entourage V from the uniformity such that $H=\{r\in E\,|\,(s,r)\in V\}$. For V there exists a basic entourage $U\equiv\{(r,t)\in E\times E\,|\,p_{f_k}(r,t)<\varepsilon,\ k=1,\ldots,m\}\subset V$. Consider the open set $G\equiv\bigcap f_k^{-1}(]f_k(s)-\varepsilon,\ f_k(s)+\varepsilon[)$ containing s. Then $G\subset H$. This means that T_0 is finer than T_1. The proposition is proved.

2.4. Almost semicontinuous functions.

Now we shall connect functions from $C^*(E, T, \varepsilon)$ with some functions on the space T.

Remind that a real-valued function x on the space T is called lower semicontinuous if for any real number a the set $x^{-1}(]a, +\infty[)$ is open. A function x on T is called *semicontinuous* if there exist lower semicontinuous functions y and z such that $x=y-z$ (see [1]—[6]). The set of all bounded semicontinuous functions on T will be denoted by $S^*(T)$. This family is a lattice algebra but not uniformly complete in general ([6]).

Let $\overline{S^*(T)}$ denote the uniform completion of $S^*(T)$. The extension $\overline{S^*(T)}$ of the family $C^*(T)$ will be called the *Sierpiňskian extension* of $C^*(T)$.

Now we shall give a functional description of this extension. A real-valued function x on T will be called *almost semicontinuous* if for every pair of intervals $[a', b']\subset]a, b[$ there exists a set $A\in\mathscr{A}(T)$ such that $x^{-1}([a', b'])\subset A\subset x^{-1}(]a, b[)$. The set of all bounded almost semicontinuous functions will be denoted by $A^*(T)$.

The family of all step-functions $x\equiv\sum a_k\chi(A_k)$ for some real numbers a_k and some sets $A_k\in\mathscr{A}(T)$ will be denoted by $S(T, \mathscr{A}(T))$. It is clear that this family is contained in the family $S^*(T)$.

Proposition 2. *The Sierpiňskian extension $\overline{S^*(T)}$ coincides with the family $A^*(T)$. The family $A^*(T)$ is a lattice algebra containing the family $S(T, \mathscr{A}(T))$ as a uniformly dense lattice subalgebra.*

Proof. Denote $A^*(T)$ by X, $S(T, \mathscr{A}(T))$ by Y and $S^*(T)$ by Z. Divide an interval containing the range of a function $x\in X$ by points a_j so that $a_{j+1}-a_j=u_n/3$. Then there exist sets A_j such that $x^{-1}([a_j, a_{j+1}])\subset A_j\subset x^{-1}(]a_{j-1}, a_{j+2}[)$. Consider the step-function $y_n\equiv\sup\{a_{j-1}\chi(A_j)|j\}\in Y$. It is easily verified that $0\leq x(t)-y_n(t)\leq u_n$ for any $t\in T$. This means that Y is uniformly dense in X.

Conversely, let for a bounded function x on T and any n there exist step-functions $y_n\in Y$ such that $|x(t)-y_n(t)|\leq v_n$ for all t and for some sequence of real numbers v_n decreasing to zero. Then for a pair of intervals $[a', b']\subset]a, b[$ there exists a number n such that $2v_n<(a'-a)\wedge(b-b')$. Consider the Alexandrov set $A\equiv y_n^{-1}(]a+v_n, b-v_n[)$. Then $x^{-1}([a', b'])\subset A\subset x^{-1}(]a, b[)$. Thus $x\in X$.

It follows from these two properties that X is a uniformly complete lattice algebra, containing Y as a uniformly dense lattice subalgebra.

Now let z be a lower semicontinuous bounded function. Take for a pair of intervals $[a', b']\subset]a, b[$ the number $c\equiv(b'+b)/2$. Consider the Alexandrov set $A\equiv z^{-1}(]a, c[)$. Then $z^{-1}([a', b'])\subset A\subset z^{-1}(]a, b[)$. Therefore z is an almost semicontinuous function. As a result we conclude that $Z\subset X$. This has as a consequence that X coincides with the Sierpiňskian extension. The proposition is proved.

Further let E be a perfect saturated preimage of T lifting separable covering.

Real-valued functions f and g on E will be called *equivalent* if for any n there exists a cohesive covering $\{C_k, \varepsilon^{-1}A_k\}$ of E such that $|f(s)-g(s)|<u_n$ for any $s\notin\bigcup(\varepsilon^{-1}A_k\setminus C_k)$. In this case we shall write $f\sim g$. It follows from Lemma 10 that this relation is indeed an equivalence relation.

Lemma 15. *The following assertions are equivalent for a function $f\in C^*(E)$:*
a) $f\in C^*(E, T, \varepsilon)$;
b) *there is a (unique) function $x\in A^*(T)$ such that $f\sim x\circ\varepsilon$.*

PROOF. Denote $A^*(T)$ by X, $S(T, \mathscr{A}(T))$ by Y and $C^*(E, T, \varepsilon)$ by Φ.

a)$\Rightarrow$b) Let $f\in\Phi$. It was established in the proof of Lemma 14 that there exist cohesive coverings $\{C_j, \varepsilon^{-1}A_j\}$ and $\{C_k, \varepsilon^{-1}A_k\}$ and step-functions $y_m\equiv\sum a_j\chi(A_j)$ and $y_n\equiv\sum a_k\chi(A_k)$ from Y such that $|f(s)-y_m\circ\varepsilon(s)|<u_m$ for any $s\notin P_m\equiv\bigcup(\varepsilon^{-1}A_j\setminus C_j)$ and $|f(s)-y_n\circ\varepsilon(s)|<u_n$ for any $s\notin P_n\equiv\bigcup(\varepsilon^{-1}A_k\setminus C_k)$. Consider the cohesive covering $\{C_j\cap C_k, \varepsilon^{-1}(A_j\cap A_k)\}$. Let $t\in T$. Then $t\in A_j\cap A_k$ for some indexes. Therefore $E_t\subset\varepsilon^{-1}(A_j\cap A_k)$. Consequently, there exists a point $s\in E_t\cap\cap\varepsilon^{-1}(A_j\cap A_k)\cap(C_j\cap C_k)$. Then $\varepsilon s=t$. Besides $s\notin P_m\cup P_n$. From these facts we obtain $|y_m(t)-y_n(t)|\leq|y_m(t)-f(s)|+|f(s)-y_n(t)|<2u_m$ for all $n\geq m$. By virtue of the previous proposition there exists a function $x\in X$ such that $|x(t)-y_n(t)|<3u_n$ for any t. Since $|f(s)-x\circ\varepsilon(s)|<4u_n$ for any $s\notin P_n$ we have $f\sim x\circ\varepsilon$.
Assume that there exists another function $x'\in X$ satisfying to this condition. Then there exists a cohesive covering $\{C_i, \varepsilon^{-1}A_i\}$ such that $|f(s)-x'\circ\varepsilon(s)|<u_n$ for any $s\notin R_n\equiv\bigcup(\varepsilon^{-1}A_i\setminus C_i)$. Let $t\in T$. Then $t\in C_k\cap C_i$ for some indices. As it was established above in this case there exists a point $s\in\varepsilon^{-1}t\cap\varepsilon^{-1}(A_k\cap A_i)\cap\cap(C_k\cap C_j)$ such that $s\notin P_n\cup R_n$. Therefore $|x(t)-x'(t)|\leq|x(t)-y_n(t)|+|y_n(t)-f(s)|+|f(s)-x'(t)|<5u_n$. Consequently $x=x'$.

b)$\Rightarrow$a) Let for the function f there is a function $x\in X$ such that $f\sim x\circ\varepsilon$. Then for a given n there exists a cohesive covering $\{C_n, \varepsilon^{-1}A_n\}$ such that $|f(s)-x\circ\varepsilon(s)|<u_n$ for any $s\notin P_n\equiv\bigcup(\varepsilon^{-1}A_k\setminus C_k)$. By virtue of the previous proposition there exists a step-function $y_n\equiv\sum a_j\chi(A_j)\in Y$ with a partition $\{A_j\}$ such that $|x(t)-y_n(t)|<u_n$ for any t. Consider the new cohesive covering $\{C_{kj}, \varepsilon^{-1}(A_k\cap A_j)\}$ where $C_{kj}\equiv C_k$. Let $r, s\in C_{kj}\cap\varepsilon^{-1}(A_k\cap A_j)$. Then $r, s\notin P_n$. Therefore $|f(r)-f(s)|<4u_n$. According to Lemma 13 there exists a cohesive covering $\{D_{kj}, \varepsilon^{-1}(A_k\cap A_j)\}$ such that $\omega(f, D_{kj})<12u_n$. This means $f\in\Phi$. The lemma is proved.

2.5. Characterizations of Alexandrovian cover.

Now with the help of the results of the preceding sections we can give characterizations of the Alexandrovian cover. Further uniqueness is understood up to isomorphism in the category of the perfect preimages of T lifting separable covering.

THEOREM 1. *Let E be the Alexandrovian cover of T. Then*
(1) *E is the unique largest of all the perfect saturated Alexandrov determined preimages of T lifting separable covering;*
(2) *E is the unique smallest of all the perfect filled lower extremally disconnected lower disjoined preimages of T lifting separable covering and, moreover, E is the unique universal (in the sense of Bourbaki) among all such preimages;*
(3) *E is the unique perfect saturated Alexandrov determined lower extremally disconnected lower disjoined preimage of T.*

PROOF. Let $\{\hat{E}, \hat{\varepsilon}\colon \hat{E}\to T, T_S\mapsto\hat{E}_S\}$ be a preimage of T having the properties from (1). We shall denote in the proof $C^*(\hat{E}, T, \hat{\varepsilon})$ by $\hat{\Phi}$ and $A^*(T)$ by X.
By virtue of Lemma 15 for any $f\in\hat{\Phi}$ there exists a unique $x\in X$ such that $f\sim x\circ\hat{\varepsilon}$. Hence we can define correctly the mapping $k\colon\hat{\Phi}\to X$ such that $kf\equiv x$. It can be checked that k is a unit preserving vector-lattice homomorphism. Let $f>0$. Then by virtue of the saturatedness there exists a set $\hat{E}_S$, on which f takes values larger than some number $a>0$. Take n such that $u_n<a$. Then there exists a cohe-

sive covering $\{C_k, \hat{\varepsilon}^{-1}A_n\}$ such that $|f(s)-x\circ\hat{\varepsilon}(s)|<u_n$ for all $s\notin P\equiv\bigcup(\hat{\varepsilon}^{-1}A_k\setminus C_k)$. Assume that $x=0$. Then $\hat{E}_S\subset P$. But $A_k\cap S\neq\emptyset$ for some index. Therefore there exists a closed set $\emptyset\neq R\subset A_k\cap S$. Then $\hat{E}_R\subset\hat{\varepsilon}^{-1}A_k\cap\hat{E}_S$ implies $\hat{E}_R\subset\hat{\varepsilon}^{-1}A_k\setminus C_k$ but this is impossible. Therefore $x>0$. Consequently, k is injective.

Verify that $\mathrm{cl}\,\hat{\varepsilon}\,\mathrm{coz}\,f=\mathrm{cl}\,\mathrm{coz}\,kf$. Denote the left-hand set by P and the right-hand set by Q. Let $s\in C\equiv\mathrm{coz}\,f$ and $t\equiv\hat{\varepsilon}s\notin Q$. Then there exists a function $g\in C^*(T)$ such that $t\in\mathrm{coz}\,g\subset T\setminus Q$. From here $s\in C_1\equiv\mathrm{coz}\,(g\circ\hat{\varepsilon})$. As $C\cap C_1\neq\emptyset$ then $kf\wedge g=k(f\wedge g\circ\hat{\varepsilon})\neq\emptyset$ but this is impossible. Consequently, $\hat{\varepsilon}C\subset Q$. Assume that there exists a function $g\in C^*(T)$ with the cozero set G such that $G\cap P=\emptyset$ and $G\cap Q\neq\emptyset$. Consider the function $f_1\equiv g\circ\hat{\varepsilon}$. As $k(f\wedge f_1)=kf\wedge g\neq 0$ then $f\wedge f_1\neq 0$, but this is false.

Now let $\{E, \varepsilon\colon E\to T, T_S\mapsto E_S\}$ be a preimage of T with the properties from 2.

Assume that for some Alexandrov set A there exists a set $U\in\triangle(E)$ such that for any S there exists a representation $S=(\bigcup S_j)\cup(\bigcup S_k)$ for some sequences of compact sets S_j and S_k such that $\bigcup E_j\subset U\cap\varepsilon^{-1}A\subset U\cup\varepsilon^{-1}A\subset E\setminus\bigcup E_k$ where $E_r\equiv E_{S_r}$ for $r\in\{j,k\}$. Then the set U with this property is unique. In fact let for the set A there exist another set $U'\in\triangle(E)$ such that for any S there exists a representation $S=\bigcup S_p\cup\bigcup S_q$ such that $\bigcup E_p\subset U'\cap\varepsilon^{-1}A\subset U'\cup\varepsilon^{-1}A\subset E\setminus\bigcup E_q$ where $E_r\equiv E_{S_r}$ for $r\in\{p,q\}$. Consider the sets $S_{jp}\equiv S_j\cap S_p$. Similarly define the sets S_{jq}, S_{kp} and S_{kp}. Denote $E_{S_{rs}}$ by E_{rs} for any $r,s\in\{j,k,p,q\}$. Then $E_{rs}\subset E_r\cup E_s$. Therefore $(\bigcup E_{jp})\cup(\bigcup E_{jq})\subset U\cup\varepsilon^{-1}A\subset U\cup\varepsilon^{-1}A\subset E\setminus(\bigcup E_{kp})\cup(\bigcup E_{kq})$ and similarly $(\bigcup E_{jp})\cup(\bigcup E_{kp})\subset U'\cap\varepsilon^{-1}A\subset U'\cup\varepsilon^{-1}A\subset E\setminus(\bigcup E_{jq}(\bigcup(\bigcup E_{kq})$. Consequently, $U\triangle U'\subset(U\triangle\varepsilon^{-1}A)\cup(U'\triangle\varepsilon^{-1}A)\subset E\setminus((\bigcup E_{jp})\cup(\bigcup E_{jq})\cup(\bigcup E_{kp})\cup(\bigcup E_{kq}))$. As the preimage E is filled we obtained $(U\triangle U')\cap E_S=\emptyset$. Since this property is valid for any S we have $U=U'$.

Denote this U by iA. Prove that iA is defined for all Alexandrov sets A. Let G be an open set. Then $G\cap S=\bigcup S_j$ and $(T\setminus G)\cap S=\bigcup S_k$ for some sequence of compact subsets of S. Consider the open-closed set $U\equiv\mathrm{cl}\,\varepsilon^{-1}G\in\triangle(E)$. Consider the sets $E_r\equiv E_{S_r}$ for $r\in\{j,k\}$. By virtue of the lower disjoinedness of the preimage E we get $\bigcup E_j\subset U\cap\varepsilon^{-1}G\subset U\cup\varepsilon^{-1}G\subset E\setminus\bigcup E_k$. Consequently, $U=iG$. If F is a closed set then $iF=E\setminus i(T\setminus F)$. If $A=G\cap F$ then consider the sets $V\equiv iG$, $W\equiv iF$ and $U\equiv V\cup W$. We have $\bigcup E_j\subset V\cap\varepsilon^{-1}G\subset V\cup\varepsilon^{-1}G\subset E\setminus\bigcup E_k$ and $\bigcup E_p\subset W\cap\varepsilon^{-1}F\subset W\cup\varepsilon^{-1}F\subset E\setminus\bigcup E_q$ for some representations $S=\bigcup S_j\cup\bigcup S_k$ and $S=\bigcup S_p\cup\bigcup S_q$. Consider as above the sets S_{rs} and E_{rs}. From $\bigcup E_{jp}\subset$ $\subset U\cap\varepsilon^{-1}A\subset U\cup\varepsilon^{-1}A\subset E\setminus(\bigcup E_{jq})\cup(\bigcup E_{kp})\cup(\bigcup E_{kq})$ we get $U=iA$. Finally, let $A=\bigcup(G_m\cap F_m)$. Denote $G_m\cap F_m$ by A_m. As it is established just now for $U_m\equiv iA_m$ there exist representations $S=\bigcup S_j^m\cup\bigcup S_k^m$ such that $\bigcup E_j^m\subset U_m\cup\varepsilon^{-1}A_m\subset$ $\subset U_m\cup\varepsilon^{-1}A_m\subset E\setminus\bigcup E_k^m$. Applying the mapping ε to the given chain of inclusions we obtain $\bigcup S_j^m\subset A_m\subset T\setminus\bigcup S_k^m$. Therefore $\bigcup S_j^m=A_m\cap S$ and $\bigcup S_k^m=S\setminus A_m$. Consider the set $U\equiv\bigcup U_m$. Represent the set $S\setminus A$ in a form $S\setminus A=\bigcup S_r$ for some sequence of compact sets S_r. Consider the sets $S_{kr}^m\equiv S_k^m\cap S_r$. Since $\bigcup\limits_k S_{kr}^m=$ $=S_r$ the set $\bigcup\limits_k E_{kr}^m$ is dense in E_r. From the inclusion $U_m\cap\bigcup\limits_k E_{kr}^m\subset U_m\cap\bigcup\limits_k E_{kr}^m=\emptyset$ we get $U\cap E_r=\emptyset$. Besides $\varepsilon^{-1}A\cap E_r=\emptyset$. As a result we obtain $\bigcup\limits_m\bigcup E_j^m\subset U\cap\varepsilon^{-1}A\subset$ $U\cup\varepsilon^{-1}A\subset E\setminus\bigcup E_r$. In addition, $S=(\bigcup\limits_m\bigcup S_j^m)\cup(\bigcup S_r)$. Hence $U=iA$.

Thus we can consider the mapping $i\colon\mathscr{A}(T)\to\triangle(E)$. Check that this mapping is a Boolean homomorphism. Let $iA=U$ and $iA'=U'$. Then $\bigcup E_j\subset U\cap\varepsilon^{-1}A\subset\subset U\cup\varepsilon^{-1}A\subset E\setminus\bigcup E_k$ and $\bigcup E_p\subset U'\cap\varepsilon^{-1}A'\subset U'\cup\varepsilon^{-1}A'\subset E\setminus\bigcup E_q$ for some representations $S=\bigcup S_j\cup\bigcup S_k$ and $S=\bigcup S_p\cup\bigcup S_q$. From here we obtain $(\bigcup E_{jp})\cup(\bigcup E_{jq})\cup(\bigcup E_{kp})\subset(\bigcup E_j)\cup(\bigcup E_p)\subset(U\cup U')\cap\varepsilon^{-1}(A\cup A')\subset(U\cup U')\cup\cup\varepsilon^{-1}(A\cup A')\subset E\setminus((\bigcup E_k)\cap(\bigcup E_q))\subset E\setminus\bigcup E_{kq}$. This means that $i(A\cup A')=U\cup U'$. Consequently, i preserves the supremum. Let $U=iT$. Then $\bigcup E_j\subset U\subset E\subset E\setminus\bigcup E_k$ implies $U\cap E_S=E_S$ for any S. Hence $U=E$. Therefore i preserves the unit. It is evident that i preserves the complementation. If $A\neq\emptyset$ then for some point $t\in A$ consider the set $S\equiv\{t\}$. In this case $S_j\neq\emptyset$ implies $E_j\neq\emptyset$. Hence $U\neq\emptyset$. Therefore the homomorphism i is injective.

Let $x=\sum a_k\chi(A_k)\in S(T,\mathscr{A}(T))$. Consider the function $f\equiv\sum a_k\chi(iA_k)\in C^*(E)$. Since i is an injective Boolean homomorphism we can define correctly the injective vector-lattice homomorphism $r\colon S(T,\mathscr{A}(T))\to C^*(E)$ by setting $rx\equiv f$. By virtue of Proposition 2 we can extend it up to an injective vector-lattice homomorphism $r\colon X\to C^*(E)$. It can be checked that $rf=f\circ\varepsilon$ for any function $f\in C^*(T)$.

Consider the injective unit preserving vector-lattice homomorphism $v\colon\hat{\Phi}\to C^*(E)$ such that $v\equiv r\circ k$. Let $t\in E$. Consider the sets $\Gamma\equiv\{f\in\hat{\Phi}|t\in\operatorname{coz}vf\}$ and $P\equiv\hat{\varepsilon}^{-1}\varepsilon t$. Assume that for $f\in\Gamma$ $P\cap\operatorname{cl}\operatorname{coz}f=\emptyset$. Then $\varepsilon t\in G\equiv T\setminus\operatorname{cl}\hat{\varepsilon}\operatorname{coz}f$ implies $t\in\varepsilon^{-1}G\subset iG$. Let $x\equiv\chi(G)$. Then $rx=\chi(iG)$. As $x\wedge kf=0$ then $rx\wedge vf=0$ but this is impossible. Hence $P\cap\operatorname{cl}\operatorname{coz}f\neq\emptyset$ for any function $f\in\Gamma$. Let $f_1,f_2\in\Gamma$. Then $f_1\wedge f_2\in\Gamma$ implies $P\cap\operatorname{cl}\operatorname{coz}f_1\cap\operatorname{cl}\operatorname{coz}f_2\neq\emptyset$. Consequently, $P_t\equiv\bigcap\{P\cap\cap\operatorname{cl}\operatorname{coz}f|f\in\Gamma\}\neq\emptyset$ by virtue of the compactness of P. Assume that there exist $s_1,s_2\in P_t$. Then there exists a function $f_1\in\hat{\Phi}$ such that $0\le f\le\mathbf{1}$, $s_1\in\operatorname{int}\{s\in\hat{E}|f_1(s)=0\}$ and $s_2\in\operatorname{int}\{s\in E|f_1(s)=1\}$. Consider the function $f_2\equiv\mathbf{1}-f_1$. Then $vf_1+vf_2=\mathbf{1}$. Assume that $t\in\operatorname{coz}vf_1$. Then $s_1\in\operatorname{cl}\operatorname{coz}f_1$ but this is false. Hence $t\in\operatorname{coz}vf_2$ implies $s_2\in\operatorname{cl}\operatorname{coz}f_2$ but this is false, too. Consequently, P_t consists of only one point. Therefore we can define correctly the mapping $\gamma\colon E\to\hat{E}$ by setting $\gamma t\equiv P_t$. This mapping is continuous. In fact, let G be a neighbourhood of the point $s\equiv\gamma t$. Consider the function $g_1\in\hat{\Phi}$ such that $0\le g_1\le\mathbf{1}$, $s\in\operatorname{int}\{s\in\hat{E}|g_1(s)=1\}$ and $\operatorname{cl}\operatorname{coz}g_1\subset G$. Let $g_2\equiv\mathbf{1}-g_1$. Assuming $t\in\operatorname{coz}vg_2$ we conclude $s\in\operatorname{cl}\operatorname{coz}g_2$ but this is false. Hence $t\in U\equiv\operatorname{coz}vg_1$. Let $t_1\in U$. Then $\gamma t_1\in\operatorname{cl}\operatorname{coz}g_1\subset G$. This gives us the continuity of γ.

This mapping is surjective. In fact, consider a point $s\in\hat{E}$, a positive real number a, the mapping $u\colon\hat{\Phi}\to\hat{\Phi}$ such that $uf\equiv(f-a\mathbf{1})\vee 0$, the set $\Gamma\equiv\{0\le f\in\hat{\Phi}|s\in\operatorname{coz}uf\}$ and the set $P\equiv\varepsilon^{-1}\hat{\varepsilon}s$. Assume that the set $P\cap\operatorname{cl}\operatorname{coz}vuf$ is empty for an $f\in\Gamma$. Then there exists a function $g\in C^*(T)$ such that $\hat{\varepsilon}s\in\operatorname{coz}g\subset T\setminus\varepsilon\operatorname{cl}\operatorname{coz}vuf$. Consider the function $f_1\equiv g\circ\hat{\varepsilon}$. Then $s\in\operatorname{coz}f_1$ means that $f_1\wedge uf\neq 0$. On the other hand $\operatorname{coz}(g\circ\varepsilon)\cap\operatorname{cl}\operatorname{coz}vuf=\emptyset$ implies $v(f_1\wedge uf)=rg\wedge vuf=0$. From here $f_1\wedge uf=0$. It follows from the given contradiction that the mentioned set is not empty. Let $f_1,f_2\in\Gamma$. Since $uf_1\wedge uf_2=u(f_1\wedge f_2)$ then $f_1\wedge f_2\in\Gamma$. Therefore $P\cap\operatorname{cl}\operatorname{coz}vuf_1\cap\cap\operatorname{cl}\operatorname{coz}vuf_2\supset P\cap\operatorname{cl}\operatorname{coz}(vu(f_1\wedge f_2))\neq\emptyset$. Then it follows from the compactness of P that there exists a point $t\in\bigcap\{P\cap\operatorname{cl}\operatorname{coz}vuf|f\in\Gamma\}$. But $\operatorname{cl}\operatorname{coz}vuf\subset\operatorname{coz}vf$. Therefore $t\in\bigcap\{\operatorname{coz}vf|f\in\Gamma\}$. Consequently, $\gamma t\in\bigcap\{\operatorname{cl}\operatorname{coz}f|f\in\Gamma\}=s$.

It follows from the definition of the mapping γ that $\hat{\varepsilon}\circ\gamma=\varepsilon$. This has as a consequence that the mapping γ is perfect ([8], VI, § 2, 56).

Check that $\gamma E_S=\hat{E}_S$. Assume that there exists a function $0\le f\in\hat{\Phi}$ with the cozero set C such that $C\cap\hat{E}_S\neq\emptyset$ and $\operatorname{cl}C\cap\gamma E_S=\emptyset$. Assume that there exists a point $t\in\operatorname{coz}vf\cap E_S$. Then $\gamma t\in\operatorname{cl}C\cap\gamma E_S=\emptyset$ but this is impossible. Therefore $\operatorname{coz}vf\cap E_S=\emptyset$. Consider the function $x\equiv kf$. Assume that there exists a point $t\in S$

such that $x(t)\equiv a>0$. Then there exists an Alexandrov set A such that $t\in \in x^{-1}([a/2, a])\subset A\subset x^{-1}(]a/3, +\infty[)$. Consider the function $y\equiv\alpha\chi(A)/3\leqq x$. From the property $\operatorname{coz} ry\cap E_S=\emptyset$ there follows $iA\cap E_S=\emptyset$. By the construction of the mapping i there exists a representation $S=\cup S_j\cup\cup S_k$ such that $\cup E_j\subset A\cap\varepsilon^{-1}A\subset \subset iA\cup\varepsilon^{-1}A\subset E\setminus\cup E_k$. Consequently, $\cup E_j=\emptyset$ implies $A\cap S=\emptyset$ but this is false. Hence $S\cap\operatorname{coz} x=\emptyset$. Consider the sets $U_n\equiv f^{-1}(]u_n, +\infty[)$. Then there exist a number n and a set $R\subset S$ such that $\hat{E}_R\subset U_n\cap\hat{E}_S$. Since $f\sim x\circ\hat{\varepsilon}$ there exists a cohesive covering $\{C_k, \hat{\varepsilon}^{-1}A_k\}$ such that $|f(s)-x\circ\hat{\varepsilon}(s)|<u_n$ for all $s\notin P_n\equiv \equiv\cup(\hat{\varepsilon}^{-1}A_k\setminus C_k)$. We can suppose that the sets A_k are mutually disjoint. As $\hat{E}_S\cap \hat{E}_S\cap\operatorname{coz}(x\circ\hat{\varepsilon})=\emptyset$ then $\hat{E}_R\subset P_n$. Further for some k there exists a non-empty compact set $R_1\subset A_k\cap R$. As a result we obtain $\hat{E}_{R_1}\subset\hat{\varepsilon}^{-1}A_k\cap P_n=\hat{\varepsilon}^{-1}A_k\setminus C_k$. However, this is impossible. From this contradiction we conclude that $\hat{E}_S\subset\gamma E_S$.

Conversely, assume that there exists a function $0\leqq f\in\Phi$ with the cozero set C such that $C\cap\gamma E_S\neq\emptyset$ and $C\cap\hat{E}_S=\emptyset$. Consider the function $x\equiv kf$. Assume that there exists a point $t\in S$ such that $x(t)\equiv a>0$. Consider, as above, the set A. Take the number n such that $u_n<a/3$. Then there exists a cohesive covering $\{C_k, \hat{\varepsilon}^{-1}A_k\}$ such that $|f(s)-x\circ\hat{\varepsilon}(s)|<u_n$ for all $s\notin P_n\equiv\cup(\hat{\varepsilon}^{-1}A_k\setminus C_k)$. We can suppose that $\{A_k\}$ is a partition. We have $\hat{\varepsilon}^{-1}A\cap\hat{E}_S\subset P_n$. As $A\cap S\neq\emptyset$ then $S_k\equiv A_k\cap A\cap S\neq\emptyset$ for some k. Therefore there exists a non-empty compact set $R\subset S_k$. From here we get $\hat{E}_R\subset\hat{\varepsilon}^{-1}A_k\cap\hat{\varepsilon}^{-1}A\cap\hat{E}_S\subset\hat{\varepsilon}^{-1}A_k\cap P_n=\hat{\varepsilon}^{-1}A_k\setminus C_k$. Since this is impossible, $S\cap\operatorname{coz} x=\emptyset$. By Proposition 2 for x there exist step-functions $x_n\equiv\sum a_{np}\chi(A_{np})$ with partitions $\{A_{np}\}$ such that $0\leqq x-x_n\leqq u_n\mathbb{1}$. Fix numbers n and p. By the construction of the mapping i there exists a representation $S=\cup S_j\cup\cup S_k$ such that $\cup E_j\subset iA_{np}\cap\varepsilon^{-1}A_{np}\subset iA_{np}\cup\varepsilon^{-1}A_{np}\subset E\setminus\cup E_k$. As $\cup S_j\subset S\cap A_{np}=\emptyset$ then $S=\cup S_k$. As far as $\cup E_k$ is dense in E_S we have $iA_{np}\cap E_S=\emptyset$. Consequently, $\operatorname{coz} rx_n\cap E_S=\emptyset$. But $0\leqq rx-rx_n\leqq u_n\mathbb{1}$ gives as a result $\operatorname{coz} rx\cap E_S=\emptyset$. Let $t\in\gamma^{-1}C$. Then $f(\gamma t)\equiv \equiv 2a>0$. Consider the function $f_1\equiv a\mathbb{1}-(f\wedge a\mathbb{1})$ and its cozero set C_1. In the assumption $t\in\operatorname{coz} vf_1$ we conclude $\gamma t\in\operatorname{cl} C_1$ but this is false. Therefore $vf+vf_1\geqq a\mathbb{1}$ implies $t\in\operatorname{coz} vf$. Hence $\gamma^{-1}C\cap E_S=\emptyset$. As this contradicts to the initial assumption, then indeed $\hat{E}_S=\gamma E_S$.

Thus E is larger than $\hat{E}$. Now let E be the Alexandrovian cover of T. As E has the properties from 1) and 2) simultaneously we get as a result that the Alexandrovian cover is the largest of all the preimages with the properties from 1) and the smallest of all the preimages with the properties from 2).

Let $\hat{E}$ be some other largest preimage of T. Then there are mappings $\gamma\colon E\to\hat{E}$ and $\delta\colon\hat{E}\to E$ such that E is larger than $\hat{E}$ relative to γ and $\hat{E}$ is larger than E relative to δ. Let $t\in E_S$. By virtue of the saturatedness of E we have $t=\cap\{E_R\}$. This implies $\delta\gamma t\in\cap\{E_R\}=t$. As $\cup E_S$ is dense we conclude that $\delta\circ\gamma=\mathrm{id}$. This means that γ and δ are mutually inverse homeomorphisms and so the preimages E and $\hat{E}$ are isomorphic.

The uniqueness of the smallest preimage and assertion 3) are checked in a similar manner. The theorem is proved.

This theorem will be used further for the proof of the following Theorems 2 and 3.

With the help of the previous theorem we can give the following functional characterization of the Alexandrovian cover. Remind that a family Φ of functions on E is called completely regular if the set $\{\operatorname{coz} f\,|\,f\in\Phi\}$ constitutes a base of the topology on E.

LEMMA 16. *Let E be the Alexandrovian cover of T. Then*

a) *for any function* $x \in A^*(T)$ *there exists a unique function* $f \in C^*(E, T, \varepsilon)$ *such that* $x \circ \varepsilon \sim f$;

b) *the mapping* $r: x \mapsto f$ *is a bijection between* $A^*(T)$ *and* $C^*(E, T, \varepsilon)$.

PROOF. We shall use the notations from the proof of the previous lemma. Let $x \in X$. Then $|x(t) - y_n(t)| < u_n$ for all t and some sequence $y_n \in Y$. Let $y_n = \sum a_k \chi(A_k)$, $\{A_k\}$ is a partition and $U_k \equiv iA_k$. Consider the continuous function $f_n \equiv \sum a_k \chi(U_k)$. Since $\{U_k, \varepsilon^{-1}A_k\}$ is a cohesive covering then $f_n \in \Phi$. Since $\{f_n\}$ is a Cauchy sequence there exists a function $f \in \Phi$ such that $|f(s) - f_n(s)| < 3u_n$. From here for any $s \notin P_n \equiv \bigcup(\varepsilon^{-1}A_k \setminus U_k)$ we have $|x \circ \varepsilon(s) - f(s)| \leqq |x \circ \varepsilon(s) - y_n \circ \varepsilon(s) + |f_n(s) - f(s)| < 4u_n$. Consequently, $f \sim x \circ \varepsilon$.

Assume that there exists another function f' possessing the same property. Then there exists a cohesive covering $\{C_j, \varepsilon^{-1}A_j\}$ such that $|f'(s) - x \circ \varepsilon(s)| < u_n$ for all $s \notin Q_n \equiv \bigcup(\varepsilon^{-1}A_j \setminus C_j)$. Consider the cohesive covering $\{C_j \cap U_k, \varepsilon^{-1}(A_j \cap A_k)\}$. Let $s \in H \equiv \bigcup(\varepsilon^{-1}(A_j \cap A_k) \cap (C_j \cap U_k))$. Then $s \notin P_n \cup Q_n$ implies $|f(s) - f'(s)| < 5u_n$. From the density of the set H we conclude now that the given inequality is valid for all points $s \in E$. Therefore $f = f'$. Hence the mapping r is defined correctly. It follows from the previous lemma that this mapping is bijective. The lemma is proved.

PROPOSITION 3. *Let E be the Alexandrovian cover of T. Then*

a) $\{E, \varepsilon: E \to T, T_S \mapsto E_S\}$ *is a perfect saturated preimage of T lifting separable covering;*

b) *there is a bijection* $r: x \mapsto f$ *between* $A^*(T)$ *and the completely regular family* $C^*(E, T, \varepsilon)$ *such that* $x \circ \varepsilon \sim f$;

c) *E as a preimage of T lifting separable covering is completely determined (up to isomorphism) by the properties* a)—b).

PROOF. Let $\{\hat{E}, \hat{\varepsilon}: \hat{E} \to T, T_S \mapsto \hat{E}_S\}$ be a preimage of T with the properties from a) and b). Denote $C^*(\hat{E}, T, \hat{\varepsilon})$ by $\hat{\Phi}$ and $A^*(T)$ by X. Let $\hat{r}x = f$, $\hat{r}y = g$ and $\hat{r}(x \vee y) = h$. Then for a fixed number n there exist cohesive coverings $\{C_i, \varepsilon^{-1}A_i\}$, $\{C_j, \varepsilon^{-1}A_j\}$ and $\{C_k, \varepsilon^{-1}A_k\}$ with partitions such that $|f(s) - x \circ \varepsilon(s)| < u_n$ for all $s \notin P \equiv \bigcup(\varepsilon^{-1}A_i \setminus C_i)$, $|g(s) - y \circ \varepsilon(s)| < u_n$ for all $s \notin Q \equiv \bigcup(\varepsilon^{-1}A_j \setminus C_j)$ and $|h(s) - (x + y) \circ \varepsilon(s)| < u_n$ for all $s \notin R \equiv \bigcup(\varepsilon^{-1}A_k \setminus C_k)$. Consider the cohesive covering which is the intersection of the given ones. Let $s \in H \equiv \bigcup(\varepsilon^{-1}(A_i \cap A_j \cap A_k) \cap (C_i \cap C_j \cap C_k))$. Then $s \notin P \cup Q \cup R$ implies $|h(s) - (f \vee g)(s)| < 3u_n$. From the density of H and the arbitrariness of n we conclude that $h = f \vee g$. Thus $\hat{r}$ is an isomorphism of the vector lattices.

By virtue of Proposition 1 $\hat{E}$ is Alexandrov determined. Let G be an open set from T and $x \equiv \chi(G)$. Then $x = \sup\{f_\xi \in C^*(T) | f_\xi \leqq x\}$ in X. Consider the function $f \equiv \hat{r}x$. Then $f = \sup\{\hat{r}f_\xi\} = \sup\{f_\xi \circ \hat{\varepsilon}\}$ in $\hat{\Phi}$. Hence $f = \chi(U)$ where $U \equiv \mathrm{cl}\, \hat{\varepsilon}^{-1}G$. Consequently, $U \in \Delta(\hat{E})$. This means that $\hat{E}$ is lower extremally disconnected. Let $\hat{\varepsilon}^{-1}G \cap \hat{E}_S = \emptyset$. Assume $U \cap \hat{E}_S \neq \emptyset$. By virtue of the saturatedness there exists a set $R \subset S$ such that $\hat{E}_R \subset U \cap \hat{E}_S \subset U \setminus \hat{\varepsilon}^{-1}G$. Besides there exists a cohesive covering $\{C_k, \hat{\varepsilon}^{-1}A_k\}$ with a partition $\{A_k\}$ such that $|f(s) - x \circ \varepsilon(s)| < u_k$ for all $s \notin P \equiv \bigcup(\hat{\varepsilon}^{-1}A_k \setminus C_k)$. Consequently, $\hat{E}_R \cup P$. As $R \cap A_k \neq \emptyset$ for some k there exists a compact set $R_1 \subset R \cap A_k$. Therefore $\hat{E}_{R_1} \subset \hat{\varepsilon}^{-1}A_k \cap \hat{E}_R \subset \hat{\varepsilon}^{-1}A_k \setminus C_k$ but this is impossible. From this contradiction we conclude that $\hat{E}$ is lower disjoined. On the strength of Theorem 1 we conclude that the preimages E and $\hat{E}$ are isomorphic. The proposition is proved.

This proposition also will be used in the sequel. Of course we can substitute in this proposition the set $A^*(T)$ by the uniformly dense subset $S^*(T)$.

2.6. Remarks on σ-Alexandrovian cover.

Substituting open sets by cozero sets in all the above definitions we can define the notion of the *lower σ-extremally disconnectedness*, the notion of the *lower σ-disjoinedness*, the field $\mathscr{A}_\sigma(T)$ of all *σ-Alexandrov subsets of T*, the *σ-Alexandrovian cover E_σ of T*, the notion of the *σ-Alexandrov determinedness*, the family $S^*_\sigma(T)$ of all bounded *σ-semicontinuous functions* and the family $A^*_\sigma(T)$ of all bounded *σ-almost-semicontinuous functions*. In such manner we shall obtain the very similar theory of the σ-Alexandrovian cover and its characterizations of exactly the same form.

In fact a more general form of these definitions is obtained by considering from the beginning a completely normal Alexandrov space ([16]) rather than the families of closed sets and of zero sets in a completely regular space.

§ 3. Sierpińskian and σ-Sierpińskian extensions as vector lattices

Let T be a completely regular space and $\overline{S^*(T)}$ be the uniform completion of the vector lattice $S^*(T)$ of all bounded semicontinuous functions on T. In Section 2.4 it was shown that $\overline{S^*(T)}$ coincides with the vector lattice $A^*(T)$ of all bounded almost semicontinuous functions on T. Let $u\colon C^*(T)\to A^*(T)$ be the canonical imbedding. For a countable set S consider the ideal $A^*_S(T)\equiv\{x\in A^*(T)\mid x(S)=0\}$. Then $\{A^*(T), u\colon C^*(T)\to A^*(T),\ C^*_S(T)\mapsto A^*_S(T)\}$ is a vector-lattice extension of $C^*(T)$ inheriting separable decomposition. This extension will be called the *Sierpińskian extension of $C^*(T)$*.

3.1. Functional description of Sierpińskian extension by functions on Alexandrovian cover.

Let E be the Alexandrovian cover of T and $\varepsilon\colon E\to T$ be the canonical mapping. Let $\Phi\equiv C^*(E, T, \varepsilon)$ be the vector lattice of functions on E defined in Section 2.3. Consider the injective vector-lattice homomorphism $\varphi\colon C^*(T)\to\Phi$ such that $\varphi f\equiv f\circ\varepsilon$. For a countable set S consider the ideal $\Phi_S\equiv\{f\in\Phi\mid f(E_S)=0\}$. Then $\{\Phi, \varphi\colon C^*(T)\to\Phi,\ C^*_S(T)\mapsto\Phi_S\}$ is a vector-lattice extension of $C^*(T)$ inheriting separable decomposition.

Now let $\{X, u\colon C^*(T)\to X,\ C^*_S(T)\mapsto X_S\}$ be a vector-lattice extension of $C^*(T)$ inheriting separable decomposition. Identify $C^*(T)$ with its image in X.

An element $x\in X$ will be called a *d-supremum of a set* $\{x_\xi\}\subset X$ if $x\geqq x_\xi$ and for any X_S it is valid that $\bar{x}=\sup\bar{x}_\xi$ in X/X_S. In this case we shall write $x=d-\sup x_\xi$. In a similar way a *d-infimum* of $\{x_\xi\}$ is defined.

Consider the sets $S_l(C^*(T), X)\equiv\{x\in X\mid\exists f_\xi\in C^*(T)(x=d-\sup f_\xi)\}$, $S(C^*(T), X)\equiv\{x-y\mid x, y\in S_l(C^*(T), X)\}$ and $A(C^*(T), X)\equiv\{x\in X\mid\forall n\exists x_n\in S(C^*(T), X)(|x-x_n|\leqq u_n\mathbf{1})\}$ where $u_n\equiv 1/n$.

The extension X will be called *Sierpiński generated* if $X=A(C^*(T), X)$.

For the extensions $A^*(T)$ and Φ consider the mapping $r: A^*(T) \to \Phi$ from Proposition 3 of Section 2.5.

PROPOSITION 4. *With respect to the mapping r the extensions $A^*(T)$ and Φ are isomorphic saturated Sierpiński generated lower Dedekind complete lower component extensions of $C^*(T)$ inheriting separable decomposition.*

PROOF. Denote $A^*(T)$ by X and $A_S^*(T)$ by X_S. It can be verified that $r\circ u=\varphi$ and r is an isomorphism of the vector lattices. Let $x\in X$ and $f\equiv rx$. Let $x\in X_S$ and assume that $f\notin \Phi_S$. By virtue of the saturatedness there exist a number n and a compact set $R\subset S$ such that $E_R\subset C\cap E_S\neq\emptyset$ where $C\equiv f^{-1}(]u_n, +\infty[)$. Now there exists a cohesive covering $\{C_k, \varepsilon^{-1}A_k\}$ with a partition $\{A_k\}$ such that $|f(s)-x\circ\varepsilon(s)|<u_n$ for any $s\notin P\equiv\bigcup(\varepsilon^{-1}A_k\setminus C_k)$. As $(x\circ\varepsilon)(E_R)=0$ we have $E_R\cup P$. Then $R_1\equiv R\cap A_k\neq\emptyset$ for some index implies $E_{R_1}\subset\varepsilon^{-1}A_k\setminus C_k$ but this is impossible. Thus our assumption is false. Conversely, let $f\in\Phi_S$ and assume that $x\notin X_S$. Then there exists a number n such that $S\cap B\neq\emptyset$ where $B\equiv x^{-1}(]u_n, +\infty[)$. Therefore there exist an index k and a non-empty compact set R such that $R\subset S\cap \cap B\cap A_k$. As a result we obtain $E_R\subset\varepsilon^{-1}A_k\cap P=\varepsilon^{-1}A_k\setminus C_k$ but this is impossible. Consequently, $x\in X_S$. Thus the extensions X and Φ are isomorphic.

Let Y be a proper component of X and $Y^d\not\subset X_S$. Consider the non-empty set $P\equiv\{t\in T|\forall y\in Y(y(t)=0)\}$. Then $Y=\{x\in X|x(P)=0\}$. In assumption $P\cap S=\emptyset$ from the inclusion $\bigcup\{\operatorname{coz} y|y\in Y^d\}\subset P$ we obtain $Y^d\subset X_S$ but this is false. Consequently, $R\equiv P\cap S\neq\emptyset$. Therefore $X_S\cup Y\subset X_R$. This means that X is saturated.

Now verify that X is Sierpiński generated. Let $x\in S_l^*(T)$. Then $x(t)= =\sup\{f_\xi(t)\}$ for some family $f_\xi\in C^*(T)$. It can be checked that $\bar{x}=\sup\bar{f}_\xi$ in any X/X_S. Therefore $x=d-\sup f_\xi$. Thus $S^*(T)\subset S(C^*(T), X)$. Now it follows from Proposition 2 that $X\subset A(C^*(T), X)$. As the rest of the properties of X are well-known the proposition is proved.

3.2. Characterization of Sierpinskian extension as a vector lattice.

Further uniqueness is understood up to isomorphism in the category of the vector-lattice extensions of $C^*(T)$ inheriting separable decomposition.

THEOREM 2. 1) *$A^*(T)$ is the unique largest of all the saturated Sierpiński generated extensions of $C^*(T)$ inheriting separable decomposition;*

2) *$A^*(T)$ is the unique smallest of all the σ-filled lower Dedekind complete lower component extensions of $C^*(T)$ inheriting separable decomposition and moreover $A^*(T)$ is the unique universal among all such extensions;*

3) *$A^*(T)$ is the unique saturated Sierpiński generated lower Dedekind complete lower component extension of $C^*(T)$ inheriting separable decomposition.*

PROOF. Let $\{X, u: C^*(T)\to X, C_S^*(T)\mapsto X_S\}$ be an extension having the properties from 1). On the strength of Yosida's theorem ([14]) there is a unique compact E_0 such that the vector lattice X is isomorphic to the vector lattice $C(E_0)$ relative to an isomorphism r_0. Then the mapping u generates a unique surjective continuous mapping ε_0: $E_0\to\beta T$ such that $r_0uf=f'\circ\varepsilon_0$, where f' denotes the extension of a function $f\in C^*(T)$ on βT.

Consider the space $E\equiv\varepsilon_0^{-1}T$ and the perfect mapping ε: $E\to T$ which is the restriction of ε_0. Consider the vector lattice Φ consisting of the restrictions on E of all function from $C(E_0)$, the homomorphism $r\colon X\to\Phi$ such that $rx\equiv r_0x|E$, and the homomorphism $\varphi\colon C^*(T)\to\Phi$ such that $\varphi f\equiv f^0\varepsilon$.

For a countable set S consider the ideals $\Phi_{0S}\equiv r_0X_S$ and $\Phi_S\equiv rX_S$ and the closed subsets $E_{0S}\equiv\{s\in E_0|\forall f\in\Phi_{0S}(f(s)=0)\}\neq\emptyset$ and $E_S\equiv E_{0S}\cap E$. Then $\cup E_{0S}$ is dense in E_0 and $\varepsilon_0E_{0S}=\mathrm{cl}\,T_S$. It is clear that $S_1\subset S_2$ implies $E_{S_1}\subset E_{S_2}$. Let R be a compact subset of S. Then $E_R=E_{0R}$. It follows from this fact that $E_S\neq\emptyset$ for any S.

Let $f\geqq 0$ be a function from $C(E_0)$ such that $f(E_{0S})=0$. Consider the functions $f_k\equiv(f-u_k\mathbf{1})\vee 0$. From the property $E_{0S}\cap\mathrm{cl}\,\mathrm{coz}\,f_k=\emptyset$ we conclude that $f_k\in\Phi_{0S}$. This implies that f also belongs to this set. Thus $\Phi_{0S}=\{f\in C(E_0)|f(E_{0S})=0\}$.

Let C be the cozero set of a function $f\in C(E_0)$ such that $C\cap E_{0S}\neq\emptyset$. Represent S in the form $S=\cup R_k$ for some compact subsets R_k. As X is filled we have $f\notin\Phi_{0R_k}$ for some k. Therefore $C\cap E_S\cap C\cap E_{0R_k}\neq\emptyset$. This means that E_S is dense in E_{0S}. As a consequence we get $\Phi_S=\{f\in\Phi|f(E_{0S})=0\}$ and $\varepsilon E_S=T_S$.

Besides we established that E is dense in E_0. Hence the triplet $\{\Phi,\varphi\colon C^*(T)\to\Phi, C_S^*(T)\mapsto\Phi_S\}$ is an extension isomorphic to the initial one.

In addition we get that $\cup E_S$ is dense in E. Consequently, E is the preimage of T lifting separable covering.

Let G be an open set in E and $G\cap E_S\neq\emptyset$. Take a non-empty regular closed set $F\subset G$ such that $\mathrm{int}\,F\cap E_S\neq\emptyset$. Consider the proper component $Y\equiv\{f\in\Phi|f(F)=0\}$. As $Y^d\not\subset\Phi_S$ we get by virtue of the saturatedness that there exists an ideal Φ_R containing the set $\Phi_S\cup Y$. This means that $E_R\subset E_S\cap G$. Thus E is a saturated preimage.

Now let f be an element from $S_l(C^*(T),\Phi)$. Then $f=d-\sup\{f_\xi\circ\varepsilon\}$. Consider the lower semicontinuous function x on T such that $x(t)\equiv\sup\{f_\xi(t)\}$. Divide an interval containing the ranges of f and x by points a_j so that $a_{j+1}-a_j=u_n/4$. Consider the cozero sets $C_j\equiv f^{-1}(]a_{j-2},a_{j+3}[)$. Further there exist Alexandrov sets A_j such that $x^{-1}([a_j,a_{j+1}])\subset A_j\subset x^{-1}(]a_{j-1},a_{j+2}[)$. Assume that there exists a set $E_S\subset\varepsilon^{-1}A_j\setminus C_j$. Fix a point $s\in E_S$ and take the point $t\equiv\varepsilon s$. If $f(s)\leqq a_{j-2}$ then $f_\xi(t)\leqq a_{j-2}$ for any ξ implies $x(t)\leqq a_{j-2}$. As $t\in A_j$ we have $x(t)>a_{j-1}$. From this contradiction we conclude that $f(s)\geqq a_{j+3}$. Consider the number $a\equiv(a_{j+2}+$ $+a_{j+3})/2$ and the open neighbourhood $U\equiv\{p\in E|f(p)>a\}$ of s. Take a function $g\in\Phi$ such that $s\in\mathrm{coz}\,g\subset U$ and $0<g\leqq b\mathbf{1}$, where $b\equiv(a_{j+3}-a_{j+2})/2$, and consider the function $h\equiv f-g$. By virtue of the saturatedness there exists a set $E_R\subset E_S\cap\mathrm{coz}\,q$. Then for any $p\in E_R$ we have $(f_\xi\circ\varepsilon(p)-h(p))\vee 0\leqq(a_{j+2}-a+b)\vee 0=0$. Therefore $\overline{f_\xi\circ\varepsilon}\leqq\bar{h}<\bar{f}$ in Φ/Φ_R for any ξ but this contradicts to the definition of d-supremum. Thus $\{C_j,\varepsilon^{-1}A_j\}$ is a cohesive covering. Since $|f(s)-x\circ\varepsilon(s)|<u_n$ for any $s\notin\cup$ $\cup(\varepsilon^{-1}A_j\setminus C_j)$ we obtain $f\sim x\circ\varepsilon$. According to Lemma 15 we conclude that $f\in C^*(E,T,\varepsilon)$. Hence $S(C^*(T),\Phi)\subset C^*(E,T,\varepsilon)$ and finally $\Phi\subset C^*(E,T,\varepsilon)$. This implies the complete regularity of the latter family. By virtue of Proposition 1 we deduce that the preimage E is Alexandrov determined.

Now let $\{\hat{X},\hat{u}\colon C^*(T)\to\hat{X}, C_S^*(T)\to\hat{X}_S\}$ be an extension having the properties from 2). Consider as it was done above the isomorphic extension $\{\hat{\Phi},\hat{\varphi}\colon C^*(T)\to$ $\to\hat{\Phi}, C_S^*(T)\to\hat{\Phi}_S\}$ for the corresponding preimage $\{\hat{E},\hat{\varepsilon}\colon\hat{E}\to T, T_S\mapsto\hat{E}_S\}$.

Let an $S=\cup S_K$ for some sequence of subsets S_K. Then $\hat{\Phi}_S=\cap\hat{\Phi}_{S_K}$ implies that $\cup\hat{E}_{S_K}$ is dense in $\hat{E}_S$. This means that the preimage $\hat{E}$ is σ-filled. Let G be an open set from T. Consider the family $\{f_\xi\}$ consisting of all continuous functions which are smaller than the characteristic function of the set G. Consider the function $f \equiv \sup\{\hat{\varphi} f_\xi\}$. Then $f(s)=1$ for any $s \in \hat{\varepsilon}^{-1} G$ and $f(s)=0$ for any $s \notin U \equiv \operatorname{cl} \hat{\varepsilon}^{-1} G$. From the continuity of the function f we conclude that $f=x(U)$. Hence U is open-closed. Thus the preimage $\hat{E}$ is lower extremally disconnected.

Let $\hat{\varepsilon}^{-1} G \cap \hat{E}_S = \emptyset$. Then $f_\xi \in C_S^*(T)$ implies $f \in \hat{\Phi}_S$. Therefore $U \cap \hat{E}_S = \emptyset$. This means that $\hat{E}$ is lower disjoined.

On the strength of Theorem 1 there exists a mapping $\gamma\colon \hat{E} \to E$ such that $\hat{E}$ is larger than E relative to γ.

Let $f \in S_l(C^*(T), \Phi)$. Then $f = d - \sup \varphi f_\xi$. Therefore there exists $g \equiv \sup \hat{\varphi} f_\xi$ because $\hat{\Phi}$ is lower Dedekind complete. Check that $f \circ \gamma = g$. Fix a set S. Assume that there exists a point $s \in \hat{E}_S$ such that $(f \circ \gamma)(s) \neq g(s)$. Consider the point $t \equiv \gamma s$ and the number $a \equiv (f(t) + g(s))/2$. If $f(t) < g(s)$ then there exists an open neighbourhood G of s such that $(f \circ \gamma)(p) < b < a < c < g(p)$ for some b and c and for any $p \in G$. Assume that $(\hat{\varphi} f_\xi)(p) \leqq a$ for any $p \in G$ and any ξ. Take a function $u \in \hat{\Phi}$ such that $s \in \operatorname{coz} u \subset G$ and $0 < u \leqq (c-a)\mathbf{1}$. Consider the function $h \equiv g - u < g$. If $p \in \operatorname{coz} u$ then $h(p) > a \geqq (\hat{\varphi} f_\xi)(p)$. Therefore $h \geqq \hat{\varphi} f_\xi$ for any index implies $h \geqq g$. From this contradiction we conclude that there exist ξ and $p \in G$ such that $(\hat{\varphi} f_\xi)(p) > a$. Then $a < (\varphi f_\xi)(\gamma p) \leqq (f \circ \gamma)(p) < a$ but this is impossible. From the obtained contradiction we conclude that $g(s) < a < f(t)$. Consider the non-empty open sets $G \equiv \{p \in \hat{E} \mid g(p) < a\}$ and $U \equiv \{q \in E \mid a < f(q)\}$. Let $S = \{s_k \mid k\}$. As $\hat{E}$ is σ-filled $G \cap \gamma^{-1} U \cap \hat{E}_{s_k} \neq \emptyset$ for some k. Take a point p from the latter set. Then $\gamma p \in U \cap E_{s_k}$. By virtue of the saturatedness we have $E_{s_k} \subset U$. Assume that $f_\xi(s_k) \leqq a$ for any Greek index. Then $\varphi f_\xi(E_{s_k}) \subset]-\infty, a]$ implies $\bar{f} \leqq a\bar{1}$ in Φ/Φ_{s_k}. Hence $f(Es_k) \subset]-\infty, a]$ but this is impossible. Thus there exists an index ξ such that $f_\xi(s_k) > a$. From here we obtain $g(\hat{E}_{s_k}) \subset]a, +\infty[$. This gives us $a < g(p) < a$. From this contradiction we conclude that our initial assumption is not valid, i.e. $(f \circ \gamma)(s) = g(s)$ for any $s \in \hat{E}_S$. As the set S was taken arbitrarily this equality is valid for any $s \in \hat{E}$. So $f \circ \gamma \in \hat{\Phi}$.

As a consequence we obtain $f \circ \gamma \in \hat{\Phi}$ for any $f \in \Phi$. Therefore we can define correctly the injective vector-lattice homomorphism $v\colon \Phi \to \hat{\Phi}$ by setting $vf \equiv f \circ \gamma$. Then $\hat{\varphi} = v \circ \varphi$. Let $f \in \Phi_S$. Then $(vf)(\hat{E}_S) = 0$ implies $vf \in \hat{\Phi}_S$. Thus the extension $\hat{\Phi}$ is larger than the extension Φ. This fact is valid for the initial extensions $\hat{X}$ and X, too.

Now let Φ be the extension from Proposition 4 isomorphic to the Sierpińskian extension $A^*(T)$. As Φ has the properties from 1) and 2) simultaneously we get as a result that Φ is the largest of all the extensions with the properties from 1) and the smallest of all the extensions with the properties from 2).

Let $\hat{X}$ be some other largest extension of $C^*(T)$. Consider as it was done above the isomorphic extension $\{\hat{\Phi}, \hat{\varphi}\colon C^*(T) \to \hat{\Phi}, C_S^*(T) \mapsto \hat{\Phi}_S\}$ for the preimage $\{\hat{E}, \hat{\varepsilon}\colon \hat{E} \to T, T_S \mapsto \hat{E}_S\}$. Take some mapping $w\colon \Phi \to \hat{\Phi}$ such that $\hat{\Phi}$ is larger than Φ relative to w. Define the surjective perfect mapping $\delta\colon \hat{E} \to E$ by setting $\delta s \equiv \bigcap \{\operatorname{cl} \operatorname{coz} f \cap \varepsilon^{-1} \hat{\varepsilon} s \mid s \in \operatorname{coz} wf\}$. Then $\varepsilon \circ \delta = \hat{\varepsilon}$. Check that $wf = f \circ \delta$ for any function $0 \leqq f \in \Phi$. Assume that there exists a point s such that $(wf)(s) \neq (f \circ \delta)(s)$. If $(wf)(s) > (f \circ \delta)(s)$ then we shall consider the function $g \equiv f$ otherwise $g \equiv -f$. Denote the number $((wg)(s) + (g \circ \delta)(s))/2$ by a. Consider the function $h \equiv (g - a\mathbf{1}) \vee 0$. Take a neighbourhood G of s such that $(wg)(t) > a$ for any $t \in G$. Also take a neighbourhood U of the point δs such that $g(r) < a$ for any $r \in U$. Then $U \subset E \setminus \operatorname{coz} h$ and $G \cap \delta^{-1} U \subset \operatorname{coz} wh$. Therefore $\delta s \notin \operatorname{cl} \operatorname{coz} h$ and $\delta s \in \operatorname{cl} \operatorname{coz} h$ but this is impossible. From this contradiction we conclude that such point s does not exist.

Check that $\delta \hat{E}_S \subset E_S$. Assume that there exists a point $s \in \delta \hat{E}_S \setminus E_S$. Take a function $f\in\Phi_S$ such that $s\in\operatorname{coz}f$. Then for some point $t\in\hat{E}_S$ such that $s=\delta t$ we get $(wf)(t)\neq 0$. But on the other hand $wf\in\Phi_S$ implies $(wf)(t)=0$. It follows from this contradiction that this inclusion is valid.

Now take the mapping $v\colon \hat{\Phi}\to\Phi$ defined above. Let s be a point from E_S. In virtue of the saturatedness of the Alexandrovian cover we have $s=\bigcap\{E_R\}$. Then $\delta\gamma s\in\bigcap\{E_R\}=s$. From this fact we conclude that $\delta\gamma s=s$ for any point $s\in E$. Therefore $(wvf)(s)=f(s)$. Thus v and w are mutually inverse isomorphisms of the vector lattices. Therefore the extensions Φ and $\hat{\Phi}$ are isomorphic.

The uniqueness of the smallest extension and assertion 3) are checked in a similar way. The theorem is proved.

3.3. Remarks about σ-Sierpińskian extension.

Substituting open sets by cozero sets in all the above definitions we can define the family $S^*_\sigma(T)$ of all bounded *σ-semicontinuous functions*, the *σ-Sierpińskian extension* $S^*_\sigma(T)$, the family $A^*_\sigma(T)$ of all bounded *σ-almost semicontinuous functions*, the notion of the *σ-Sierpiński generatedness*, the notion of the *lower σ-Dedekind completeness* and the notion of the *lower σ-componentness*. In such manner we shall obtain the very similar theory of the σ-Sierpińskian extension and its characterization of exactly the same form.

§ 4. Sierpińskian and σ-Sierpińskian extensions as C-rings

Let T be a completely regular space and $\overline{S^*(T)}$ be the uniform completion of the ring $S^*(T)$ of all bounded semicontinuous functions on T. In Section 2.4 it was shown that $\overline{S^*(T)}$ coincides with the C-ring $A^*(T)$ of all bounded almost semicontinuous functions on T. Let $u\colon C^*(T)\to A^*(T)$ be the canonical imbedding. For a countable set S consider the C-ideal $A^*_S(T)\equiv\{x\in A^*(T)|x(S)=0\}$. Then $\{A^*(T), u\colon C^*(T)\to A^*(T), C^*_S(T)\mapsto A^*_S(T)\}$ is a C-ring extension of $C^*(T)$ inheriting separable decomposition. This extension will be called the *Sierpińskian extension of $C^*(T)$*.

4.1. Functional description of Sierpińskian extension by functions on Alexandrovian cover.

Let E be the Alexandrovian cover of T and $\varepsilon\colon E\to T$ be the canonical mapping. Let $\Phi\equiv C^*(E,T,\varepsilon)$ be the C-ring of functions on E defined in Section 2.3. Consider the injective ring homomorphism $\varphi\colon C^*(T)\to\Phi$ such that $\varphi f\equiv f\circ\varepsilon$. For a countable set S consider the C-ideal $\Phi_S\equiv\{f\in\Phi|f(E_S)=0\}$. Then $\{\Phi,\varphi\colon C^*(T)\to\Phi, C^*_S(T)\mapsto\Phi_S\}$ is a C-ring extension of $C^*(T)$ inheriting separable decomposition.

Remind that according to Section 1.3.1 any C-ring X is a lattice ring with respect to the order, defined by the cone $P\equiv\{x\in X|\exists y\in X(x=y^2)\}$.

LEMMA 17. a) *If $u\colon X\to Y$ is a ring homomorphism between C-rings X and Y then u is a lattice-ring homomorphism between the lattice rings X and Y.*

b) *If Y is a C-ideal in a C-ring X then Y is a lattice-ring ideal, i.e. $y\in Y$ and $|x|\leq|y|$ implies $x\in Y$.*

PROOF. The first assertion is checked in a routine way. On the strenght of Delfosse's theorem X is isomorphic to a lattice ring $C(K)$ for some compact space K.

Consider the elements $x_n \equiv \left[\left(x-\frac{1}{n}\mathbf{1}\right)\vee 0\right]+\left[\left(x+\frac{1}{n}\mathbf{1}\right)\wedge 0\right]$. Define the functions $z_n\colon K\to[-1,1]$ such that $z_n(s)\equiv x_n(s)/y(s)$ for any $s\in C\equiv \operatorname{coz} y$ and $z_n(s)\equiv 0$ for any $s\notin C$. It is clear that $H_n\equiv \operatorname{cl}\operatorname{coz} x_n\subset C$. Let an $s\in \operatorname{cl} C\setminus C$. Fix an arbitrary number ε. Take a neighbourhood G of s such that $G\cap H_n=\emptyset$. Then $z_n(G)=\{0\}\subset\subset]-\varepsilon,\varepsilon[$. This means that z_n is continuous in s. Thus $z_n\in C(K)$. As a result we obtain $x_n=z_ny\in Y$. As Y is C-ideal then Y is uniformly closed. Therefore the inequality $|x-x_n|\leq \frac{1}{n}\mathbf{1}$ implies $x\in Y$. The lemma is proved.

Now let $\{X, u\colon C^*(T)\to X, C^*_S(T)\mapsto X_S\}$ be a C-ring extension of $C^*(T)$ inheriting separable decomposition. Identify $C^*(T)$ with its image in X. By virtue of the previous lemma we can define the set $A(C^*(T), X)$ so as it has been done in Section 3.1. The extension X will be called Sierpiński generated if $X=A(C^*(T), X)$.

For the extensions $A^*(T)$ and Φ consider the mapping $r\colon A^*(T)\to\Phi$ from Proposition 3 of Section 2.5.

Proposition 5. *With respect to the mapping r the extensions $A^*(T)$ and Φ are isomorphic saturated Sierpiński generated lower continuing lower segment extensions of $C^*(T)$ inheriting separable decomposition.*

Proof. Denote $A^*(T)$ by X and $A^*_S(T)$ by X_S. It can be verified that $r\circ u=\varphi$ and r is isomorphism of the lattice rings. It has been checked in the proof of Proposition 4 that $x\in X_S$ iff $rx\in\Phi_S$. Therefore the extensions X and Φ are isomorphic.

In just the same way as in the proof of Proposition 4 it is checked that X is saturated and Sierpiński generated.

Let Y be a ring ideal in the ring $C^*(T)$ and $g\in \operatorname{Hom}^*_{C^*(T)}(y, C^*(T)\cap Y^{**})$. Let $y_1, y_2\in Y$ and $t\in \operatorname{coz} y_1\cap \operatorname{coz} y_2$. Then $(gy_1)(t)/y_1(t)=(gy_2)(t)/y_2(t)$. Consequently, we can define correctly the almost semicontinuous function $z\in X$ by setting $z(t)\equiv(gy)(t)/y(t)$ for any $y\in Y$ and any $t\in\operatorname{coz} y$ and $z(t)\equiv 0$ for any $t\notin G\equiv\equiv\bigcup\{\operatorname{coz} y\,|\,y\in Y\}$.

As $z\in Y^{**}$ we can define correctly the homomorphism $h\in\operatorname{Hom}^*_X(X, Y^{**})$ by setting $hx\equiv xz$. Let $y\in Y$ and $t\in G$. Then there exists an element $y_1\in Y$ such that $t\in\operatorname{coz} y_1$. Therefore $y_1(t)(hy)(t)=y(t)(gy_1)(t)=y_1(t)(gy)(t)$ implies $(hy)(t)==(gy)(t)$. As hy and gy belong to Y^{**} we have $(hy)(t)=0=(gy)(t)$ for any $t\notin G$. This means that $hy=gy$. Thus X is lower continuing.

Now let g and h be the homomorphisms from the definition of the lower segment and $gY\subset X_S$. Let $x\in X$ and $t\in S\cap G$. Then $t\in\operatorname{coz} y$ for some $y\in Y$ implies $y(t)(hx)(t)=x(t)(gy)(r)=0$ and hence $(hx)(t)=0$. If $t\in S\setminus G$ then $(hx)(t)=0$ because of $hx\in Y^{**}$. Consequently, $hx\in X_S$. This means that X_S is a lower segment of X. The proposition is proved.

4.2. Characterization of Sierpiński extension as a C-ring.

Further uniqueness is understood up to isomorphism in the category of the C-ring extensions of $C^*(T)$ inheriting separable decomposition.

Theorem 3. 1) *$A^*(T)$ is the unique largest of all the saturated Sierpiński generated extensions of $C^*(T)$ inheriting separable decomposition;*

2) *$A^*(T)$ is the unique smallest of all the σ-filled lower continuing lower segment*

extensions of $C^*(T)$ inheriting separable decomposition and moreover $A^*(T)$ is the unique universal among all such extensions;

3) *$A^*(T)$ is the unique saturated Sierpiński generated lower continuing lower segment extension of $C^*(T)$ inheriting separable decomposition.*

PROOF. Let $\{X, u\colon C^*(T)\to X,\ C_S^*(T)\mapsto X_S\}$ be an extension having the properties from 1). On the strength of Delfosse's theorem ([15]) there is a compact E_0 such that the lattice ring X is isomorphic to the lattice ring $C(E_0)$ relative to an isomorphism r_0. The lattice-ring homomorphism u generates a unique surjective continuous mapping $\varepsilon_0\colon E_0\to\beta T$ such that $r_0uf=f'\circ\varepsilon_0$, where f' denotes the extension of a function $f\in C^*(T)$ on βT. Further, by completely the same arguments as in the proof of Theorem 2 we obtain the preimage $\{E, \varepsilon\colon E\to T,\ T_S\mapsto E_S\}$ of T and the corresponding extension $\{\Phi, \varphi\colon C^*(T)\to\Phi,\ C_S^*(T)\mapsto\Phi_S\}$ isomorphic to the initial one.

In just the same ways as in the proof of Theorem 2 it is established that the preimage E is saturated and Alexandrov determined.

Now let $\{\hat{X}, \hat{u}\colon C^*(T)\to\hat{X}, C_S^*(T)\mapsto\hat{X}_S\}$ be an extension having the properties from 2). Consider the isomorphic extension $\{\hat{\Phi}, \hat{\varphi}\colon C^*(T)\to\hat{\Phi},\ C_S^*(T)\mapsto\hat{\Phi}_S\}$ for the corresponding preimage $\{\hat{E}, \hat{\varepsilon}\colon\hat{E}\to T, T_S\mapsto\hat{E}_S\}$. Then the preimage $\hat{E}$ is σ-filled.

Let G be an open set from T. Denote the set $\hat{\varepsilon}^{-1}G$ by V. Consider the ring $R\equiv\hat{\varphi}C^*(T)$ and the ring ideal $Y\equiv\{y\in R|\operatorname{coz} y\subset V\}$ of the ring R. Define the homomorphism $g\in\operatorname{Hom}_R^*(Y, R\cap Y^{**})$ by setting $gy\equiv y$. Then there exists a bounded $\hat{\Phi}$-module homomorphism $h\colon\hat{\Phi}\to Y^{**}$ extending g. Consider the function $u\equiv h\mathbf{1}\in\hat{\Phi}$ and the set $U\equiv\operatorname{cl} V$. It is clear that $u(\hat{E}\setminus U)=0$. Let $s\in V$. Then $s\in\operatorname{coz} y$ for some $y\in Y$. Therefore $y(s)u(s)=(gy)(s)=y(s)$ implies $u(s)=1$. Since the function u is continuous we conclude that $u=\chi(U)$ and $U\in\Delta(\hat{E})$. This means that the preimage $\hat{E}$ is lower extremally disconnected.

Let $V\cap\hat{E}_S=\emptyset$. Then $gY\subset\hat{\Phi}_S$ implies $u\in\hat{\Phi}_S$. Therefore $U\cap\hat{E}_S=\emptyset$. Thus the preimage $\hat{E}$ is lower disjoined.

On the strength of Theorem 1 there exists a mapping $\gamma\colon\hat{E}\to E$ such that $\hat{E}$ is larger than E relative to γ. Consider the mappings $i\colon\mathscr{A}(T)\to\Delta(\hat{E})$ and $v\colon\Phi\to\to C^*(\hat{E})$ from the proof of Theorem 1. As it has been established $iG=\operatorname{cl}\hat{\varepsilon}^{-1}G$ for any open set G from T. Just above it was proved that $\chi(\operatorname{cl}\hat{\varepsilon}^{-1}G)\in\hat{\Phi}$. It has as a consequence that $v\Phi\subset\hat{\Phi}$. From here we conclude that $vf=v\circ\gamma$ for any $f\in\Phi$. In fact assume that $f\geqq 0$ and there exists a point s such that $(vf)(s)\neq(f\circ\gamma)(s)$. If $(vf)(s)>(f\circ\gamma)(s)$ then we shall consider the function $g\equiv f$ otherwise $g\equiv -f$. Denote the number $((vg)(s)+(g\circ\gamma)(s))/2$ by a. Consider the function $h\equiv(g-a\mathbf{1})\vee 0$. Take a neighbourhood G of s such that $(wg)(t)>a$ for any $t\in G$. Also take a neighbourhood U of the point γs such that $g(r)<a$ for any $r\in U$. Then $U\subset E\setminus\operatorname{coz} h$ and $G\cap\gamma^{-1}U\subset\operatorname{coz} vh$. From the first inclusion we obtain $\gamma s\notin\operatorname{cl}\operatorname{coz} h$. The second inclusion by the definition of γ gives us $\gamma s\in\operatorname{cl}\operatorname{coz} h$. From this contradiction we conclude that such point s does not exist.

Further the proof is led in exactly the same way as the proof of Theorem 2.

4.3. Remarks about σ-Sierpińskian extension.

Substituting open sets by cozero sets, sets of indexes by countable sets of indexes and ideals by countably generated ideals in the corresponding definitions we can define the family $S_\sigma^*(T)$ of all bounded *σ-semicontinuous functions on T*, the family

$A_\sigma^*(T)$ of all bounded *σ-almost semicontinuous functions on T*, the notion of *lower σ-continuingness*, the notion of the *lower σ-segmentness* and the notion of the *σ-Sierpińskian generatedness*. In such manner we shall obtain the very similar theory of the *σ-Sierpińskian extension* and its characterization of exactly the same form.

REFERENCES

[1] HAUSDORFF, F., Über halbstetige Funktionen und deren Verallgemeinerung, *Math. Z.* **4** (1915), 292—309.

[2] KAPLAN, S., On the second dual of the space of continuous functions, *Trans. Amer. Math. Soc.* **86** (1957), 70—90. *MR* **19**—868.

[3] KAPLAN, S., The unbounded bidual of $C(X)$, *Illinois J. Math.* **17** (1973), 619—645. *MR* **48** # 12010.

[4] SHANNON, C., The second dual of $C(X)$, *Pacific J. Math.* **72** (1977), 237—253. *MR* **58** # 12324.

[5] RÜSCHENDORF, L., Note on semicontinuous functions, *Studia Sci. Math. Hungar.* **15** (1980), 147—149. *MR* **84**: 46047.

[6] SIERPIŃSKI, W., Sur les fonctions développables en séries absolument convergentes de fonctions continues, *Fund. Math.* **2** (1921), 15—27.

[7] ZAHAROV, V. K., Some perfect preimages connected with extensions of the family of continuous functions, *Topology. Theory and Applications* (Proc. Conf. Eger, Hungary, 1983), North-Holland, 1985.

[8] Архангельский, А. В. и Пономарёв, В. И., *Основы общей топологии в задачах и упражнениях*, Наука, Москва, 1974.

[9] SIKORSKI, R., *Boolean algebras*, Ergebnisse der Mathematik und ihrer Grenzgebiete, N. F., Heft 25, Springer-Verlag, Berlin—Göttingen—Heidelberg, 1960. *MR* **23** # A3689.

[10] SEMADENI, Z., *Banach spaces of continuous functions*, Vol. 1, Monografie Matematyczne, Tom 55, Polish Scientific Publishers, Warszawa, 1971. *MR* **45** # 5730.

[11] FUCHS, L., *Partially ordered algebraic systems*, Pergamon Press, Oxford, 1963. *MR* **30** # 2090.

[12] VULIH, B. Z., *Introduction to the theory of partially ordered spaces*, Gosudarstv. Izdat. Fiz.-Mat. Lit., Moscow, 1961 (in Russian). *MR* **24** # A3494.

[13] LUŃEMBURG, W. A. J. and ZAANEN, A. Z., *Riesz spaces*, Vol. I, North-Holland Mathematical Library, North-Holland Publishing Co., Amsterdam, 1971. *MR* **58** # 23483.

[14] YOSIDA, K., On vector lattice with a unit, *Proc. Imp. Acad. Tokyo* **17** (1941), 121—124. *MR* **3**—210.

[15] DELFOSSE, J.-P., Caractérisations d'anneaux de fonctions continues, *Ann. Soc. Sci. Bruxelles Sér. I* **89** (1975), 364—368. *MR* **52** # 1312.

[16] ALEKSANDROV, A. D., Additive set-functions in abstract spaces, *Rec. Math. [Sbornik] N. S.* **8** (1940), 307—348; **9** (1941), 563—628; **13** (1943), 169—238. *MR* **2**—315; **3**—207; **6**—275.

[17] CSÁSZÁR, A., *General topology*, Akadémiai Kiadó, Budapest, 1978.

(Received January 20, 1985)

CHAIR OF GENERAL TOPOLOGY AND GEOMETRY
FACULTY OF MECHANICS AND MATHEMATICS
MOSCOW STATE UNIVERSITY
LENINSKIE GORY
SU—117 234 MOSCOW V-234
SOVIET UNION

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 119—126

ON THE DENSITY OF MULTIPLE PACKINGS AND COVERINGS OF CONVEX DISCS

U. BOLLE

1. In the following paper I try to give estimates for the density of multiple packings and coverings of lattice translates of a compact convex disc in the euclidean plane. We say that a lattice G gives a k-packing (k-covering) of the disc K if each point of the plane is in at most (at least) k translates $\mathbf{x}+\operatorname{int}(K)/\mathbf{x}+\operatorname{cl}(K)$, $\mathbf{x}\in G$.

The density $d(G, K)$ of an arrangement of translates of K is defined by

$$d(G, K) = \lim_{R\to\infty} \frac{1}{4R^2} \sum_{\mathbf{x}+K\subseteq Q_R} m(K)$$

where $Q_R=\{\mathbf{x}|\mathbf{x}=(x_1, x_2), \max\{|x_1|, |x_2|\}\leqq R\}$ and $m(K)$ denotes the area of K. It is an easy consequence that $d(G, K)=m(K)/\Delta(G)$ if $\Delta(G)$ is the determinant of G.

We will be concerned with the optimal densities given by

$$d_k(K) = \sup\{d|\exists G: d = d(G, K) \text{ and } G \text{ packs } K\ k\text{-fold}\}$$

$$D_k(K) = \inf\{d|\exists G: d = d(G, K) \text{ and } G \text{ gives a } k\text{-covering}\}.$$

The exact values of d_k and D_k are known only in a few cases (see [1]), so we try to give estimates.

For the following let K always be a compact convex disc with nonempty interior. We choose an arbitrary point from int (K). This point is the origin of all coordinate systems to be used in this paper. The boundary of K can be represented by polar coordinates $r=r(w)$. We then only consider discs whose function r satisfies the following conditions:

(i) r is 2π-periodic and continuous,

(ii) for all $w\in[0, 2\pi]$ r has bounded and continuous one-sided derivatives of the first and second order, and there are only finitely many w for which r is not twice differentiable.

For this class of discs we will prove that there is a constant $c=c(K)>0$ with $d_k(K)\geqq k-ck^{2/5}$. A similar result holds for coverings.

Further we will show that for some of these discs, especially for all polygons, even $d_k(K)\geqq k-ck^{1/3}$.

2. For the proof we need some preparations, which are contained in this section.

1980 *Mathematics Subject Classification.* Primary 52A45.
Key words and phrases. Packing and covering, convexity.

Let G be a lattice in the plane. Then we always can find a base $\{\mathbf{a}, \mathbf{b}\}$ with $\mathbf{a}=(a, 0)$, $\mathbf{b}=(g, h)$ such that

(i) $a = \min\{|\mathbf{x}||\in G\setminus\{O\}\}$,

(ii) $0 \leqq g < a$,

(iii) $h > 0$.

We will only use bases of this kind.

Let now G be a lattice with base $\{\mathbf{a}, \mathbf{b}\}$. We say that G is of order $n(\in \mathbf{N})$ with respect to K if

$$\frac{b(K)}{n+1} \leqq h < \frac{b(K)}{n}$$

where $b(K)$ is the width of K in the direction of the Y-axis (orthogonal to $\mathbf{a}$).

We denote by $K_\alpha (0\leqq\alpha\leqq 2\pi)$ the disc K rotated around the origin in the positive sense with angle α. If $X(w)=r(w)\cos(w)$, $Y(w)=r(w)\sin(w)$ is the representation of the boundary of K then we find for K_α: $X_\alpha(w)=r(w-\alpha)\cos(w)$, $Y_\alpha(w)=$ $=r(w-\alpha)\sin(w)$. Now for an arbitrary lattice G and $y\in\mathbf{R}$ we define a function S by

$$S(y, h, \alpha) = \sum_{i\in Z} m_1\big((K'_\alpha+(0, y))\cap(\mathbf{R}\times\{ih\})\big).$$

Here m_1 means the length on a line and

$$K'_\alpha = \begin{cases} K_\alpha & \text{for coverings} \\ \operatorname{int}(K_\alpha) & \text{for packings.} \end{cases}$$

There is a simple geometric interpretation for S: the lattice points of G lie on the parallel lines $\mathbf{R}\times\{ih\}$, $i\in Z$ and S represents the sum of the lengths of the segments which are cut out by $K'_\alpha+(0, y)$ from these lines.

The function S is connected to our problem by the following

LEMMA 1. *Let* $v(x, y)=\operatorname{card}\{g|g\in G$ and $(x, y)\in g+K'_\alpha\}$ *be the packing (covering) multiplicity of the point with the coordinates* (x, y) *by* K'_α. *Then we have*

$$\int_0^a v(x, y)\,dx = S(-y, h, \alpha) \quad \textit{for all} \quad y\in\mathbf{R}.$$

PROOF. Let e_{ij} be the characteristic function of $K'_\alpha+i\mathbf{a}+j\mathbf{b}$ so that

$$e_{ij}(x, y) = \begin{cases} 1 & \text{if } (x, y)-i\mathbf{a}-j\mathbf{b}\in K'_\alpha \\ 0 & \text{otherwise,} \end{cases}$$

then

$$v(x, y) = \sum_{i,j\in Z} e_{ij}(x, y)$$

and

$$\int_0^a v(x, y)\,dx = \sum_{j\in Z}\sum_{i\in Z}\int_0^a e_{ij}(x, y)\,dx.$$

Now we have

$$\int_0^a e_{ij}(x, y)\,dx = m_1\big(([0, a] \times \{y\}) \cap (K'_\alpha + i\mathbf{a} + i\mathbf{b})\big) =$$

$$= m_1\big(([-ia, -(i-1)a] \times \{-jh\}) \cap (K'_\alpha + (0, -y))\big),$$

and consequently

$$\sum_{i \in \mathbf{Z}} \int_0^a e_{ij}(x, y)\,dx = m_1\big((\mathbf{R} \times \{-jh\}) \cap (K'_\alpha + (0, -y))\big)$$

which proves the lemma.

COROLLARY. *If G packs K'_α k-fold then $S(y, h, \alpha) \leqq ka$ for all $y \in \mathbf{R}$, similar for each k-covering $S(y, h, \alpha) \geqq ka$ for all y holds.*

We now define some further quantities: If $k \in \mathbf{N}$ and if there are k-packings of order n then

$$d_k(n, K) = \sup \{d \mid \exists G, \alpha\colon (G \text{ is of order } n \text{ for } K'_\alpha \text{ and } G \text{ packs } K'_\alpha \text{ } k\text{-fold and } d(G, K'_\alpha) = d)\}.$$

$D_k(n, K)$ can be defined in the same way for all $k \in \mathbf{N}$. And finally:

$$a_n(K) = \inf_\alpha \inf_h \sup_y hS(y, h, \alpha) \qquad 0 \leqq \alpha \leqq 2\pi$$

$$A_n(K) = \sup_\alpha \sup_h \inf_y hS(y, h, \alpha) \qquad \frac{b(K'_\alpha)}{n+1} \leqq h < \frac{b(K'_\alpha)}{n}$$

$$0 \leqq y \leqq h.$$

With these definitions we have

THEOREM 1. a) *Let $b_m(K)$ be the minimal width of K and*

$$k > \frac{a_n(K)}{b_m^2(K)}(n+1)^2 + n.$$

Then $d_k(n, K)$ exists and

$$(k-n)\,m(K)/a_n(K) \leqq d_k(n, K) \leqq km(K)/a_n(K).$$

b) *For all $k \in \mathbf{N}$:*

$$(k+n)\,m(K)/A_n(K) \geqq D_k(n, K) \geqq m(K)/A_n(K).$$

PROOF. a) If G gives a k-packing of order n of K'_α then we have by the corollary of Lemma 1:

$$S(y, h, \alpha) \leqq ka \quad \text{for all} \quad y \in \mathbf{R},$$

and consequently $a_n(K) \leqq kah$. Therefore

$$d(G, K) = \frac{m(K)}{ah} \leqq km(K)/a_n(K).$$

To prove the lower bound let $\varepsilon > 0$ be an arbitrary small number.

There are $\alpha_0 \in [0, 2\pi]$ and $h_0 \in \left]\frac{b(K_{\alpha_0})}{n+1}, \frac{b(K_{\alpha_0})}{n}\right[$ with $(1+\varepsilon)a_n(K) > h_0 \max S(y, h_0, \alpha_0)$.

By the definition of S we get

$$S(y, h_0, \alpha_0) = \big(v(x, y) - z(x, y)\big)\, a + E(x, y)$$

where v denotes the earlier used multiplicity, $z \leqq n+1$ means the number of lines $\mathbf{R} \times \{ih\}$ intersected by $K'_{\alpha_0} + (x, y)$ and E is a quantity for which $0 < E \leqq 2za$. Now

$$(1+\varepsilon)\, a_n(K)/h_0 > \max_y S(y, h_0, \alpha_0) \geqq (v-z)a + E > (v-n-1)\, a$$

and therefore

$$v < (1+\varepsilon)\frac{a_n(K)}{ah_0} + n + 1.$$

If we take

$$a = (1+\varepsilon)\frac{a_n(K)}{h_0(k-n)}, \quad g = 0 \quad \text{and} \quad G_k = \langle (a, 0), (g, h_0) \rangle$$

then we have $v(x, y) \leqq k$ for all (x, y), which means that G_k gives a k-packing of K'_{α_0}. A simple calculation shows that the chosen base satisfies the postulated conditions for all sufficiently small values of ε and the lattice G_k is of course of order n with respect to K'_{α_0}; so we get

$$d_k(n, K) \geqq d(G_k, K'_{\alpha_0}) = \frac{m(K)}{ah_0} = (k-n)\frac{m(K)}{a_n(K)}\,\frac{1}{1+\varepsilon}$$

for all $\varepsilon > 0$, which is the proposition.

I omit the proof of b) for it goes exactly along the same lines.

We now have to examine $a_n(K)$, $A_n(K)$ a bit closer to get information about $d_k(K)$ and $D_k(K)$.

3. The boundary of K_α can be represented in polar coordinates in the following way:

$$X(w) = r_\alpha(w)\cos(w), \quad Y(w) = r_\alpha(w)\sin(w),$$

$$r_\alpha(w) = r(w-\alpha).$$

We denote the extremal values of Y by $\overline{Y}_\alpha = \sup_w Y_\alpha(w)$, $\underline{Y}_\alpha = \inf_w Y_\alpha(w)$.

Possibly the boundary of K_α contains at most two segments parallel to the X-axis. Then let $0 < w_1 \leqq w_2 < w_3 \leqq w_4 < 2\pi$ be the angles corresponding to the endpoints. If there are less than two segments we have $(w_1 = w_2$ and $Y_\alpha(w_1) = \overline{Y}_\alpha)$ or $(w_3 = w_4$ and $Y_\alpha(w_3) = \underline{Y}_\alpha)$. By the convexity of K the functions $Y_{\alpha,1}\colon\,]w_4 - 2\pi, w_1[\to]\underline{Y}_\alpha, \overline{Y}_\alpha[$ and $Y_{\alpha,2}\colon\,]w_2, w_3[\to]\underline{Y}_\alpha, \overline{Y}_\alpha[$ are one-to-one. Let $\lambda_i (i=1, 2)$ be the corresponding inverse functions. With these definitions we get

$$m_1\big((K'_\alpha + (0, y)) \cap (\mathbf{R}\, x\{jh\})\big) = X_\alpha\big(\lambda_1(jh-y)\big) - X_\alpha\big(\lambda_2(jh-y)\big)$$

if the intersection on the left side is not a boundary segment. So we find

$$S(y, h, \alpha) = \sum_{\underline{Y}_\alpha < jh-y < \overline{Y}_\alpha} \left(X_\alpha(\lambda_1(jh-y)) - X_\alpha(\lambda_2(jh-y))\right) + R,$$

where R is the sum of the lengths of the boundary segments. Now by the Dirichlet sum formula:

$$\sum_{\underline{Y}_\alpha < jh-y < \overline{Y}_\alpha} X_\alpha(\lambda_1(jh-y)) = \sum_{p=-\infty}^{\infty} \int_{(\underline{Y}_\alpha + y)h}^{(\overline{Y}_\alpha + y)/h} X_\alpha(\lambda_1(zh-y)) \cos(2\pi pz)\, dz =$$

$$= \frac{1}{h} \sum_{p=-\infty}^{\infty} \int_{w_4 - 2\pi}^{w_1} X_\alpha(w) Y_\alpha'(w) \cos(H_p(w))\, dw,$$

where

$$H_p(w) = \frac{2\pi p}{h}(Y_\alpha(w) + y),$$

and

$$\sum_{\underline{Y}_\alpha < jh-y < \overline{Y}_\alpha} X_\alpha(\lambda_2(jh-y)) = -\frac{1}{h} \sum_{p=-\infty}^{\infty} \int_{w_2}^{w_3} X_\alpha(w) Y_\alpha'(w) \cos(H_p(w))\, dw.$$

Therefore:

$$hS(y, h, \alpha) = \sum_{p=-\infty}^{+\infty} \int_0^{2\pi} X_\alpha(w) Y_\alpha'(w) \cos(H_p(w))\, dw + hR =$$

$$= \int_0^{2\pi} X_\alpha(w) Y_\alpha'(w)\, dw + 2 \sum_{p=1}^{\infty} \int_0^{2\pi} X_\alpha(w) Y_\alpha'(w) \cos(H_p(w))\, dw + hR.$$

In the following I use instead of hS the function T defined by $T(y, h, \alpha) = hS(y, h, \alpha) - hR$.

We have $a_n(K) = \inf_\alpha \inf_h \sup_y T(y, h, \alpha)$, $A_n(K) = \sup_\alpha \sup_h \inf_y T(Y, h, \alpha)$ for in the case of packings $T = hS$ as always $R = 0$ whereas in the case of coverings certainly $\inf_y hS(y, h, \alpha) = \inf_y T(y, h, \alpha)$. So we may write

$$T(y, h, \alpha) = \int_0^{2\pi} X_\alpha(w) Y_\alpha'(w)\, dw + 2 \sum_{p=1}^{\infty} \int_0^{2\pi} X_\alpha(w) Y_\alpha'(w) \cos(H_p(w))\, dw.$$

PROPOSITION.

$$\int_0^{2\pi} X_\alpha(w) Y_\alpha'(w)\, dw = m(K).$$

PROOF.

$$X_\alpha(w) Y_\alpha'(w) = r_\alpha \cos(w)\left(r_\alpha' \sin(w) + r_\alpha \cos(w)\right) =$$

$$= r_\alpha^2 + r_\alpha \sin(w)\left(r_\alpha' \cos(w) - r_\alpha \sin(w)\right) = X_\alpha'(w) Y_\alpha(w) + r_\alpha^2$$

for all but at most finitely many $w\in[0, 2\pi]$. Integration by parts gives

$$\int_0^{2\pi} X_\alpha(w)\, Y_\alpha'(w)\, dw = \frac{1}{2}\int_0^{2\pi} r_\alpha^2(w)\, dw = m(K).$$

We now have to estimate the remaining integrals

$$C_p = \int_0^{2\pi} X_\alpha(w)\, Y_\alpha'(w) \cos\left(H_p(w)\right) dw \quad (p\in\mathbf{N}).$$

Again integration by parts shows

$$C_p = -\frac{h}{2\pi p}\int_0^{2\pi} X_\alpha'(w) \sin\left(H_p(w)\right) dw.$$

Even if the boundary of K contains segments there are only finitely many α, for which the boundary of K_α contains segments parallel to the X-axis. Let

$$A = \{\alpha | \alpha\in[0, 2\pi] \text{ and } w_1 = w_2 \text{ and } w_3 = w_4\}$$

and, in the following, always $\alpha\in A$.

Especially simple to handle are discs with the following property: there is an $\alpha_0\in A$ with $Y_{\alpha_0}'(w_1\pm 0)\neq 0$ and $Y_{\alpha_0}'(w_3\pm 0)\neq 0$, which means that the boundary of K_{α_0} contains two "vertices" with tangent lines parallel to each other and the X-axis. Discs with this property I call „acute". All polygons are acute discs.

Lemma 2. *If K is an acute disc there are constants $c_1, c_2 > 0$ with $a_n(K)\leqq m(K)+c_1 n^{-2}$ and $A_n(K)\geqq m(K)-c_2 n^{-2}$.*

Proof. As K is acute there exists $\alpha_0\in A$ and $c_3 > 0$ so that $Y_{\alpha_0}'(w)\geqq c_3$ for all w. Now

$$C_p = -\frac{h}{2\pi p}\int_{0\#}^{2\pi} X_\alpha'(w) \sin\left(H_p(w)\right) dw = \left(\frac{h}{2\pi p}\right)^2 \int_0^{2\pi} \frac{X_{\alpha_0}'(w)}{Y_{\alpha_0}'(w)} \left(\cos\left(H_p(w)\right)\right)' dw$$

and

$$|C_p| \leqq 2\left(\frac{h}{2\pi p}\right)^2 \frac{\max\limits_w X_{\alpha_0}'(w)}{c_3}$$

by the second mean value theorem, because K is convex and therefore $X_{\alpha_0}'/Y_{\alpha_0}'$ is monotonic in $[w_3-2\pi, w_1]$ and $[w_1, w_3]$. So we get

$$T(y, h, \alpha_0) = m(K) + 2\sum_{p=1}^{\infty} C_p = m(K) + O(n^{-2}),$$

where the O-constant can be chosen independent of y and h; and

$$a_n(K) = \inf_\alpha \inf_h \sup_y T(y, h, \alpha) \leqq \inf_h \sup_y T(y, h, \alpha_0) \leqq m(K) + c_1 n^{-2},$$

$$A_n(K) = \sup_\alpha \sup_h \inf_y T(y, h, \alpha) \geqq \sup_h \inf_y T(y, h, \alpha_0) \geqq m(K) - c_2 n^{-2}.$$

Let us now consider the general case.

It becomes clear from the foregoing that the order of magnitude of C_p depends essentially on the behaviour of Y_α near the extremal values $\overline{Y}_\alpha, \underline{Y}_\alpha$, we therefore split

C_p:

$$\int_0^{2\pi} = \int_0^{w_1} + \int_{w_1}^{\pi} + \int_{\pi}^{w_3} + \int_{w_3}^{2\pi}.$$

As all integrals can be estimated in the same way we only take the first one.

We can choose $\alpha_0 \in A$ with the following property: *if* $Y'_{\alpha_0}(w_1-0)=0$ *then* $Y''_{\alpha_0}(w_1-0)\neq 0$ *and if* $Y'_{\alpha_0}(w_1+0)=0$ *then* $Y''_{\alpha_0}(w_1+0)\neq 0$ *and similar for* w_3.

PROOF. If $Y'_\alpha(w)=0$ and $Y''_\alpha(w)=0$ this means that the curvature

$$\varkappa(w) = \frac{Y''_\alpha(w)X'_\alpha(w)-X''_\alpha(w)Y'_\alpha(w)}{(X'^2_\alpha(w)+Y'^2_\alpha(w))^{3/2}}$$

of the boundary of K_α is zero at w. Now if the curvature of ∂K is zero for all points where it is defined then K is a polygon and hence acute. So we may assume that there is $\alpha\in A$ such that Y_α is twice differentiable at w_1 and $\varkappa(w_1)\neq 0$. Since Y''_α is continuous at w_1 the curvature does not vanish in a neighbourhood of w_1. If $\varkappa(w_3)\neq 0$ then we are ready. If $\varkappa(w_3)$ exists and is $=0$ then there is a w near w_3 with $\varkappa(w)\neq 0$, since $\alpha\in A$. So rotate K_α a little and we have the desired situation. If Y_α is not twice differentiable at w_3 then we rotate K_α a little and even get $Y'_\alpha(w_3\pm 0)\neq 0$ or are in one of the former cases.

If $Y'_\alpha(w_1-0)\neq 0$ then the proof of Lemma 2 shows that

$$\int_0^{w_1} = O\left(\frac{h}{p}\right).$$

So let us assume that $Y'_\alpha(w_1-0)=0$ and $Y''_\alpha(w_1-0)\neq 0$ for the following. Take $\varepsilon=\sqrt{h/p}$ and $M=\max|X'_\alpha(w)|$. Then we have

$$\int_0^{w_1} X'_\alpha(w)\sin(H_p(w))\,dw = \int_0^{w_1-\varepsilon} X'_\alpha(w)\sin(H_p(w))\,dw +$$

$$+\int_{w_1-\varepsilon}^{w_1} X'_\alpha(w)\sin(H_p(w))\,dw = I_1+I_2.$$

For the partial integrals I_1, I_2 we find

$$|I_1| = \frac{h}{2\pi p}\left|\int_0^{w_1-\varepsilon}\frac{X'_\alpha(w)}{Y'_\alpha(w)}(\cos(H_p(w)))'\,dw\right| \leqq \frac{h}{\pi p}\frac{M}{|Y'_\alpha(w_1-\varepsilon)|} =$$

$$= \frac{h}{\pi p}\frac{M}{\varepsilon|Y''_\alpha(w_1-\theta\varepsilon)|} \leqq c_4\frac{h}{p\varepsilon}$$

again by the second mean value theorem and

$$|I_2| = \left|\int_{w_1-\varepsilon}^{w_1} X'_\alpha(w)\sin(H_p(w))\,dw\right| \leqq M\varepsilon$$

and consequently

$$|I_1+I_2| \leqq c_5\sqrt{h/p},$$

where $c_5>0$ and independent of h and p. Therefore we have

$$|C_p| \leqq c_6\left(\frac{h}{p}\right)^{3/2}.$$

With essentially the same proof as for Lemma 2 we get

LEMMA 3. *There are constants* $c_7, c_8>0$ *with*

$$a_n(K) \leqq m(K)+c_7 n^{-3/2}$$

and

$$A_n(K) \geqq m(K)-c_8 n^{-3/2}.$$

As an immediate consequence of Theorem 1 we now find

THEOREM 2. *There are constants* $c_9, c_{10}>0$ *with*

$$\frac{d_k(n,K)}{k} \geqq \left(1-\frac{n}{k}\right)(1-c_9 n^{-\beta}) \quad \textit{for} \quad k > \frac{a_n(K)}{b_m^2(K)}(n+1)^2+n$$

and

$$\frac{D_k(n,K)}{k} \leqq \left(1+\frac{n}{k}\right)(1+c_{10} n^{-\beta})$$

for all k. *Here* $\beta=3/2$ *is the general case. If the disc* K *is acute then we may take* $\beta=2$.

Finally we can give estimates for $d_k(K)$ and $D_k(K)$.

THEOREM 3. *There are constants* $c_{11}, c_{12}>0$ *depending only on* K *such that*

$$\frac{d_k(K)}{k} \geqq 1-c_{11}k^{-\gamma}, \quad \frac{D_k(K)}{k} \leqq 1+c_{12}k^{-\gamma},$$

where $\gamma=\beta/(1+\beta)$, *i.e.* $\gamma=3/5$ *in the general case and* $\gamma=2/3$ *for acute discs.*

PROOF. For a given k take $n_n=\left[\frac{1}{k^{1+\beta}}\right]$. Then

$$k \geqq c_{13} n^{1+\beta} \geqq c_{14} n^{5/2}$$

and the condition of Theorem 1 is satisfied. So

$$\frac{d_k(K)}{k} \geqq \frac{d_k(n_k, K)}{k} \geqq \left(1-\frac{n_k}{k}\right)(1-c_9 n_k^{-\beta}) \geqq 1-c_{11}k^{-\gamma}.$$

The proof for coverings follows the same pattern.

REFERENCE

[1] FEJES TÓTH, G., New results in the theory of packing and covering, *Convexity and its applications*, ed. by P. M. Gruber and J. M. Wills, Birkhäuser, Basel—Boston, Mass., 1983, 318—359. *MR* **85i**: 52007.

(Received May 15, 1985)

GYMNASIUM LÜCHOW
D—3130 LÜCHOW
FEDERAL REPUBLIC OF GERMANY

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 127—132

A TAUBERIAN THEOREM FOR COMBINED LIMITS OF FUNCTIONS OF TWO VARIABLES

PETER REIMNITZ

1. Introduction

Frequently the asymptotic behaviour of the series

$$(1) \qquad \sum_{m=0}^{M} \beta_{n,m} \quad (M \to \infty)$$

with n depending on M is of interest. Solutions to (1) are often difficult to obtain.

A connected but in many cases analytically more tractable problem is the asymptotic solution for

$$(2) \qquad \sum_{m=0}^{\infty} \beta_{n,m} t^m \quad (t \to 1)$$

where n may depend on t.

In dynamic programming (1) corresponds by proper standardization to average cost. (2) corresponds to discounted cost, the terms $\beta_{n,m}$ represent cost per unit time.

A more mathematical application where results of (2) can be used to solve (1), is given in Section 3 of this article, where a generalized arcsine-law for the oscillating random walk is derived.

In this article we want to connect the two sequences in (1) and (2). For this we consider functions $f(t, z)$ of the form:

$$(3) \qquad f(t, z) = \sum_{0}^{\infty} b_m(z) t^m \quad (t \in C,\ z \in C).$$

Let

$$(4) \qquad B_a(s) = \lim_{t \to 1} (1-t)^{\beta} f(t, 1-(s(1-t))^a) \quad (\beta \in (0, \infty),\ a \in (0, \infty),\ s \in C).$$

Under certain regularity conditions to be stated later, we obtain

$$B_a(s) = \lim_{M \to \infty} M^{-\beta} \sum_{m=0}^{M} \sum_{n=0}^{\infty} b_{m,n} \Gamma(an+\beta+1) M^{-an} s^{an}$$

and

$$B_a(s) = \lim_{m \to \infty} m^{-\beta+1} \sum_{n=0}^{\infty} b_{m,n} \Gamma(an+\beta+1) M^{-an} s^{an}.$$

Research supported by "Deutsche Forschungsgemeinschaft" under SFB 72.

1980 *Mathematics Subject Classification.* Primary 40E05; Secondary 4400.

Key words and phrases. Tauberian theorem.

Further, we will show

$$\lim_{M\to\infty} M^{-\beta} \sum_{m=0}^{M} b_m\big(1-(x/M)^a\big) = L(B_a)(x), \tag{5}$$

and

$$\lim_{m\to\infty} m_m^{-\beta+1} b\big(1-(x/m)^a\big) = L'(B_a)(x) \qquad (x\in R),$$

where L and L' are specific linear operators given later.

2. Results

Throughout this article we assume:

(A) $f(t, z)$ to be bivariate analytical in $t\in C$ and $z\in C$ for $|t|<1$ and $|1-z|<\varepsilon$ $\varepsilon>0$.

(B) The function $f(t, z)$ to be such that

$$B_a(s) = \lim\,(1-t)^\beta f\big(t,\, 1-(s(1-t))^a\big)$$

exists for fixed values of a $\big(a\in(0, \infty)\big)$ and b $\big(b\in(0, \infty)\big)$ and is analytic in s^a for $|s^a|<r$, $r>0$.

Part of the results to be obtained depends on a monotonicity assumption:

(C) The sequences $(b_{m,n})_{m=0}^{\infty}$ are monotonic in m for n fixed, where $(b_{m,n})$ are defined to be the coefficients of the Taylor series expansions of the functions $b_m(z)$ given in (3).

The following lemma provides the basis for the remaining results.

Lemma. *Under the assumptions* (A) *and* (B), *we have for $B_a(s)$ as defined in* (4), *provided* $|s^a|<\min(r/2, \varepsilon/2)$:

$$B_a(s) = \sum_{n=0}^{\infty} \lim_{M\to\infty} M^{-\beta} \sum_{m=0}^{M} b_{m,n}(s/M)^{an}\Gamma(an+\beta+1) \tag{6}$$

(Γ denotes the Gamma function).

If we assume (A), (B) and (C), *we have:*

$$B_a(s) = \sum_{n=0}^{\infty} \lim_{m\to\infty} m^{-\beta+1} b_{m,n}\Gamma(an+\beta)\,(s/m)^{an}. \tag{7}$$

Proof. For $|t|<1$ and $|s^a|<\min(r/2, \varepsilon/2)$, we may write

$$f(t, z) = \sum_{m=0}^{\infty} t^m \sum_{n=0}^{\infty} b_{m,n}(1-z)^n.$$

As $|t|<1$ and $|s^a|<\min(r/2, \varepsilon/2)$ a change of summations is allowed because of absolute convergence of the infinite series involved $\big(|s^a|<\min(r/2, \varepsilon/2)$ implies $|1-(1-(s(1-t))^a|<\min(r, \varepsilon)$, for $|t|<1$), hence:

(8) $$f(t, 1-(s(1-t))^a) = \sum_{n=0}^{\infty} \{(1-t)^{an} \sum_{m=0}^{\infty} t^m b_{m,n}\} s^{an}.$$

As $|s^a|<r$ (Condition (B)), we have, by an appropriate choice of $B_{a,n}$,

(9) $$B_a(s) = \sum_{n=0}^{\infty} B_{a,n} s^{an}.$$

From (8) and (9) it follows:

(10) $$\lim_{t\to 1} (1-t)^{an+\beta} \sum_{m=0}^{\infty} t^m b_{m,n} = B_{a,n}.$$

A Tauberian theorem ([2], Theorem 5, p. 447) gives:

(11) $$\sum_{m=0}^{M} b_{m,n} \sim \Gamma(an+\beta+1)^{(-1)} M^{an+\beta} B_{a,n}$$

hence equation (6).

Under the additional monotonicity assumption (Condition (C)) the above mentioned Tauberian theorem gives:

(12) $$b_{m,n} \sim \Gamma(an+\beta)^{(-1)} m^{an+\beta-1} B_{a,n},$$

from which (7) follows. ∎

By changing limit and infinite summation, equations (4) can be derived from equations (6) and (7).

In what follows, a sufficient condition is given, under which the above mentioned change is allowed.

(D) Assume that the function $f(t, z)$ of (3) fulfils the condition: For all n, m and $M \in N$, there exists an $R \in (0, \infty)$ independent of n, m and M so that

(13) $$M^{-an-\beta} \sum_{m=0}^{M} b_{m,n} \Gamma(an+\beta+1) < R^{an}$$

(here $b_{m,n}$ are the terms of the Taylor expansion of the functions $b_m(z)$ of (3)).

THEOREM 1. *Under the assumptions* (A), (B) *and* (D), *we have*

(14) $$\lim_{M\to\infty} M^{-\beta} \sum_{m=0}^{M} \sum_{n=0}^{\infty} b_{m,n} \Gamma(an+\beta+1)(s/M)^{an} = B_a(s)$$

for $|s^a|<1/R$.

Assuming additionally (C), *we have:*

(15) $$\lim_{m\to\infty} m^{-\beta+1} \sum_{n=0}^{\infty} b_{m,n} \Gamma(an+\beta)(s/m)^{an} = B_a(s)$$

for $|s|<1/R$.

The proofs follow from our lemma and from the fact that the convergence of a sequence of analytic functions to an analytic function implies the convergence of the corresponding Taylor series coefficients to the coefficients of the Taylor series expansion of the limiting function.

Frequently, $b_m(z)$ is itself a transform, see e.g. our example in Section 3, where $b_m(z)$ is a Mellin transform. Then a direct inversion of $B_a(s)$ may be possible. In other cases one might be interested in getting explicitly:

$$\lim_{M\to\infty} M^{-\beta} \sum b_m(1-(x/M)^a)$$

or

$$\lim_{m\to\infty} m^{-\beta+1} b_m(1-(x/M)^a).$$

For these cases, inversion formulae for the generating functions $B_a(s)$ are needed. We give such inversion formulae in Theorem 2.

THEOREM 2. *Under the assumptions* (A) *and* (B), *we have:*

$$\lim_{M\to\infty} M^{-\beta} \sum b_m(1-(x/M)^a) = (2\pi i)^{-1} \int_C y^{-\beta-1} e^y B_a(x/y)\, dy. \tag{16}$$

C is the path of integration starting at $-\infty$ on the real axis, it circles the origin once in the positive direction with radius larger than $|x/r|$ (r as in assumption (B)) *and returns to $-\infty$; the initial and final argument of y to be $-\pi$ and $+\pi$, respectively. Under the assumptions* (A), (B) *and* (E), *we have:*

$$\lim_{m\to\infty} m^{-\beta+1} b_m(1-(x/m)^a) = (2\pi i)^{-1} \int_C y^{-\beta} e^y B_a(x/y)\, dy \tag{17}$$

(the path of integration C being as above).

PROOF. We show the validity of (16). Equation (17) follows analogously using (7) instead of (6).

Because of condition (B) we have for $|x|<\delta$, $\delta>0$:

$$\int_C |y^{-\beta-1} e^y B_a(x/y)|\, dy < \infty,$$

where $B_a(s)$ is defined in (6). Let

$$I(f)(x) = (2\pi i)^{-1} \int_C y^{-\beta-1} e^y f(x/y)\, dy,$$

where C is defined as above. As $|x/y|<r$, we have

$$\begin{aligned} I(B_a)(x) &= I\Big(\sum_{n=0}^{\infty} \lim_{M\to\infty} M^{-\beta} \sum_{m=0}^{M} b_{m,n} M^{-an} \Gamma(an+\beta+1)\,(x/.)^{an}\Big) \\ &= \sum_{n=0}^{\infty} I\Big(\lim_{M\to\infty} M^{-\beta} \sum_{m=0}^{M} b_{m,n} (1/.)^{an}\Big) M^{-an} \Gamma(an+\beta+1)\, x^{an} \\ &= \sum_{n=0}^{\infty} \lim_{M\to\infty} M^{-\beta} \sum_{m=0}^{M} b_{m,n} M^{-an} \Gamma(an+\beta+1)\, x^{an} \\ &\quad * (2\pi i)^{-1} \int_C y^{-\beta-1} e^y (1/y)^{an}\, dy. \end{aligned}$$

As

$$(2\pi i)^{-1} \int\limits_C y^{-z} e^y \, dy = 1/\Gamma(z)$$

(see [1], p. 187 equation (5—11)), we have

$$I(B_a)(x) = \sum_{n=0}^{\infty} \lim_{M\to\infty} M^{-\beta} \sum_{m=0}^{M} b_{m,n} (x/M)^{an},$$

hence

$$I(B_a)(x) = \lim_{M\to\infty} M^{-\beta} \sum_{m=0}^{M} \sum_{n=0}^{\infty} b_{m,n} (x/M)^{an}$$

$$= \lim_{M\to\infty} M^{-\beta} \sum_{m=0}^{M} b_m (1-(x/M)^a). \quad \blacksquare$$

REMARK. The inversion formulae do not require condition D. Hence it is possible to obtain equation (5) without the special series representation (6) or (7) of the function $B_a(s)$ to be valid.

In Section 3, we wish to demonstrate the usefulness of Theorem 1 by applying it to a problem in the theory of random walks. The result is a special case of the result in Reimnitz [4], where Lemma 1 was implicitly used.

3. An application

We are interested in finding the limiting distribution of the number of times an oscillating random walk stays positive. Here, an oscillating random walk is a Markov chain $(Z_n)_1^\infty$ defined by

$$Z_0 = z, \quad z \in \mathbf{R}$$

$$Z_{n+1} = Z_n + \begin{cases} Y_{1,n+1} & \text{if } Z_n \geq 0, \\ Y_{2,n+1} & \text{if } Z_n < 0. \end{cases}$$

$(Y_{1,n})_1^\infty$ and $(Y_{2,n})_1^\infty$ are two independent sequences of independent identically distributed random variables (see [3]).

Let $N_n = \sum_{n=0}^{\infty} 1_{(Z_i \geq 0)}$ (here, for any set A, $1_A(\omega) = 1$ if $\omega \in A$ and $1_A(\omega) = 0$ for $\omega \notin A$). We want to obtain the distribution of N_n/n as $n \to \infty$.

For the distributions of $Y_{1,n}$ and $Y_{2,n}$, we assume:

$$\hat{v}(s) = \mathsf{E}[\exp(sY_{1,1})] = \frac{v}{s+b} + \mathsf{E}[\exp(sY_{1,1}) \, 1_{(Y_{1,1}>0)}],$$

$$\hat{\mu}(s) = \mathsf{E}[\exp(sY_{2,1})] = \frac{\mu}{a-s} + \mathsf{E}[\exp(sY_{2,1}) \, 1_{(Y_{2,1}<0)}].$$

Here, $\mu, v, a, b \in (0, \infty)$. Further $\hat{\mu}(s)$ and $\hat{v}(s)$ exist in an open neighbourhood of zero and also

$$\mathsf{E}[Y_{1,1}] = \hat{\nu}'(s)_{\rfloor s=0} = \mathsf{E}[Y_{2,1}] = \hat{\mu}'(s)_{\rfloor s=0} = 0,$$

$$\mathsf{E}[Y_{1,1}^2] = \hat{\nu}''(s)_{\rfloor s=0} = \sigma_1^2 \quad \text{and} \quad \mathsf{E}[Y_{1,2}^2] = \hat{\mu}'(s)_{\rfloor s=0} = \sigma_2^2.$$

We find

$$\Phi(t, \varrho) = \sum_{n=0}^{\infty} \mathsf{E}[\varrho^{N_n}]\, t^n =$$

$$= (1-\varrho t)^{-1} + \big((1-\varrho t)^{-1} + (1-t)^{-1}\big) e^{-v(\varrho t)z} \frac{u(t)}{a} \frac{\big(b-v(\varrho t)\big)\big(v(\varrho t)+a\big)}{b\big(v(\varrho t)+u(\varrho t)\big)}$$

for $z \neq 0$, where, $u(t)$ $(v(t))$ is the root of $\hat{\mu}(u) = 1/t$ $(\hat{\nu}(v) = 1/t$, respectively) (see [3]).

One can easily verify that all assumptions made for the proof of Theorem 1 are fulfilled. Applying the theorem, we find $\lim_{t\to 1} (1-t)\,\Phi(t, \varrho^{1-t})$ to be essentially the Stieltjes transform of the distribution of $\lim_{m\to\infty} N_m/m$, more precisely:

$$\lim_{t\to 1} (1-t)\,\Phi(t, \varrho^{1-t}) = \mathsf{E}\left[\lim_{m\to\infty} \left(1 - \frac{N_m}{m} \log \varrho\right)^{-1}\right].$$

Somewhat tedious calculations give

$$\lim_{t\to 1} (1-t)\Phi(t, \varrho^{1-t}) =$$

$$= \big(1 - R\sqrt{1-\log \varrho}\big)^{-1} + R^2\big(R^2(1-\log \varrho) + R\sqrt{1-\log}\big)^{-1}.$$

Inversion of the Stieltjes transform, by taking its imaginary part, gives the density (with respect to the Lebesgue measure) of $\lim_{m\to\infty} N_m/m$ to be:

$$g(y) = \frac{R}{\pi} \big(y(1-y)\big)^{-1/2} \big(y + R^2(1-y)\big)^{-1} 1_{(0,1)}(y),$$

$$R = \sigma_2/\sigma_1.$$

Acknowledgement. The author wants to express his sincere gratitude to an unknown referee for improvements on an earlier version of this paper.

REFERENCES

[1] Carrier, G. F., Krook, M. and Pearson, C. E., *Functions of a complex variable; Theory and technique*, McGraw-Hill, New York, 1966. *MR* **36** # 5308.

[2] Feller, W., *An introduction to probability theory and its application II*, 2nd ed., J. Wiley, New York, 1971. *MR* **42** # 5292.

[3] Kemperman, J. H. B., The oscillating random walk, *Stochastic Processes Appl.* **2** (1974), 1—29. *MR* **50** # 14940.

[4] Reimnitz, P., An arcsine-law for the oscillating random walk, *Studia Sci. Math. Hungar.* **20** (1985), 439—450.

(Received January 2, 1986)

BAYER AG
PHARMA FORSCHUNGSZENTRUM
INSTITUT FÜR BIOMETRIE
APRATHER WEG
D—5600 WUPPERTAL 1
FEDERAL REPUBLIC OF GERMANY

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 133—137

ON THE FIXED POINT ALGEBRAS FOR φ-CONDITIONAL EXPECTATIONS IN VON NEUMANN ALGEBRAS

CARLO CECCHINI and DÉNES PETZ

Introduction and notations

It is a well-known fact that a conditional expectation (i.e., projection of norm one) of a von Neumann algebra M onto a von Neumann subalgebra N exists under rather restrictive conditions ([10]). Let φ be a faithful normal state on M. The φ-conditional expectation $\mathsf{E}_\varphi^{M,N}$ (or shortly E) from M into N preserving φ was studied in [1] and [8], and used in [5] and in [2]. The φ-conditional expectation exists for any triple (M, N, φ) and reduces to projection of norm one whenever this exists.

The aim of this paper is to study the fixed point subalgebras of E. This is motivated both by the interest of comparing φ-conditional expectations with conditional expectations (when N itself is a fixed point algebra) and by importance of this notion in [2] for the construction of Markov chains.

Our general reference on the modular theory of von Neumann algebras is [9]. Let the von Neumann algebra M act on a Hilbert space H. We shall use the spaces $L(p, M, \varphi)$ introduced in [3] and strongly connected (isometrically isomorphic) to the spaces defined and studied in [6] and [7]. The latter spaces, denoted by $L^p(\omega')$, consists of operators and are constructed with respect to a faithful normal semifinite weight ω' on the commutant M' of M, while the former consist of linear forms defined on the lineal $D(H, \varphi)$. If $T\in L^p(\omega')$ then $q(T)$ denotes the corresponding complex form in $L(p, M, \varphi)$, on the other hand if $q\in L(p, M, \varphi)$ then $T_p(q)$ is defined by $T_p(q(S))=S$ $(S\in L^p(\omega'))$.

If $1/p+1/s=1/r\leqq 1$ then for $q_1\in L(p, M, \varphi)$ and $q_2\in L(s, M, \varphi)$ a product

$$P_{(p,s)}[q_1, q_2] = q(T_p(q_1)T_s(q_2))\in L(r, M, \varphi)$$

is defined in [3] and further studied in [4].

Let us remind that $L(1, M, \varphi)$ is canonically isometrically isomorphic to M_*, and we denote this isomorphism by $\iota_\varphi\colon M_*\to L(1, M, \varphi)$. We emphasize that since $L(p, M, \varphi)\subset L(1, M, \varphi)$, $\iota_\varphi^{-1}(q)$ is a functional on M if $q\in L(p, M, \varphi)$ and $1\leqq p\leqq\infty$.

Results

Let M_1, M_2 and M_3 be von Neumann algebras such that $M_1\subset M_2\subset M_3\subset B(H)$. We fix a faithful normal state φ_3 on M_3 (ω_1' on M_1'). Let $\varphi_i=\varphi_3|M_i$ and $\omega_i'=\omega_1'|M_i'$ $(i=1, 2, 3)$. σ^i will stand for the modular group of φ_i $(i=1, 2, 3)$.

1980 *Mathematics Subject Classification* (1985 Revision). Primary 46L50; Secondary 46L30.
Key words and phrases. Von Neumann algebras, conditional expectations, fixed points.

THEOREM. *The following conditions are equivalent.*

(i) *The φ_3-conditional expectation $E\colon M_3 \to M_2$ leaves M_1 pointwise invariant.*

(ii) *There exists a completely positive mapping $\alpha\colon M_3 \to M_2$ such that $\alpha(I)=I$, $\varphi_3 \circ \alpha = \varphi_3$ and $\alpha(a_1)=a_1$ for all $a_1 \in M_1$.*

(iii) *There exists a von Neumann subalgebra M of M_3 such that $M_1 \subset M \subset M_2$ and there is a conditional expectation $F\colon M_3 \to M$ with the property $\varphi_3 \circ F = \varphi_3$.*

(iv) $\sigma_t^3(M_1) \subset M_2 \quad (t \in \mathbf{R})$.

(iv)′ $\sigma_t^3 | M_1 = \sigma_t^2 | M_1 \quad (t \in \mathbf{R})$.

(v) *For all* $a \in M_1$

$$\iota_3^{-1} q_3(a) | M_2 = \iota_2^{-1} q_2(a).$$

(v)′ *For all* $a \in M_2$

$$\iota_3^{-1} q_3(a) | M_1 = \iota_3^{-1} q_2(a) | M_1.$$

(vi) *For all* $a, b \in M_1$ *and* $1 \leqq p, s \leqq \infty$ *with* $1/p + 1/s \leqq 1$

$$P_{(p,s)}[q_2(a), q_2(b)](\xi) = P_{(p,s)}[q_3(a), q_3(b)](\xi)$$

if $\xi \in D(H, \varphi_3)$.

PROOF. (i)→(ii): Obvious.

(ii)→(iii): By the mean ergodic theorem $n^{-1} \sum_{i=0}^{n-1} \alpha^i$ converges in the strong operator topology to a projection F of norm one and mapping M_3 onto the fixed point algebra M of α.

(iii)→(iv): According to Takesaki's theorem ([10], [1]) $\sigma_t^3(M) \subset M$ and hence $\sigma_t^3(M_1) \subset M_2$. The equivalence (iv)↔(iv)′ is contained in [1].

(i)↔(v): Using the convention of [5] we abbreviate Connes' spatial derivative ([6]) of φ_i with respect to ω_i' as d_i. Now $\iota_3^{-1} q_3(a)$ is the same as ψ_a in [3] and similarly to the proof of 3.7 Proposition of [3] we have

$$[\iota_3^{-1} q_3(a)](b) = \int d_3^{1/2} a d_3^{1/2} b \, d\omega_3'.$$

By 2.7 Lemma in [3] we can write

$$\int d_3^{1/2} a d_3^{1/2} b \, d\omega_3' = \langle \pi_3(a) \Phi, J_3 \pi_3(b^*) \Phi \rangle$$

where π_i is the GNS representation with φ_i and Φ is the corresponding cyclic and separating vector ($i=2, 3$). (π_2 is identified with the subrepresentation of π_3.) Reference to the definition of the φ-conditional expectation ([1], cf. [8]) yields

$$\langle \pi_3(a)\Phi, J_3\pi_3(b^*)\Phi\rangle = \langle \pi_2(a)\Phi, J_2\pi_2(b^*)\Phi\rangle.$$

Starting with the right-hand side of (v) we carry out a similar argument:

$$[\iota_2^{-1}q_2(a)](b) = \int d_2^{1/2}ad_2^{1/2}bd\omega_2' = \langle \pi_2(a)\Phi, J_2\pi_2(b^*)\Phi\rangle.$$

Clearly this equals to $\langle \pi_2(E(a))\Phi, J_2\pi_2(b^*)\Phi\rangle$ for all $b\in M_2$ if and only if $a=\mathsf{E}(a)$.

(iv)$'\to$(vi): Let $\xi\in D(H, \varphi_3)$ be fixed. We recall 3.4 Theorem of [4]. The mapping G_i: $S\times S\to \mathbf{C}$ defined by

$$G_i(z, w) = \int d_i^{(1-w)/2}ad_i^{(w+z)/2}bd_i^{(1-w)/2}K_i(\xi)\,d\omega_i'$$

$(z, w\in S=\{v\in \mathbf{C}\colon\ 0\leqq \operatorname{Re} v\leqq 1\}$ and $K_i(\xi)=\pi_i^{-1}(J_i|R_i(\xi)|^2J_i))$ is bounded and continuous on $S\times S$ and analytic on $S^0=\{v\in \mathbf{C}\colon\ 0<\operatorname{Re} v<1\}$ in each of the variables whenever the other one is fixed $(i=2, 3)$. Moreover,

$$P_{(p,s)}[q_i(a), q_i(b)](\xi) = G_i(1/p, 1/s)$$

$(i=2, 3)$. To prove (vi) we check that

$$G_2(it_1, it_2) = G_3(it_1, it_2)$$

$$G_2(1+it_1, it_2) = G_3(1+it_1, it_2)$$

$$G_2(it_1, 1+it_2) = G_3(it_1, 1+it_2)$$

$$G_2(1+it_1, 1+it_2) = G_3(1+it_1, 1+it_2)$$

for all $t_1, t_2\in \mathbf{R}$.

$$\begin{aligned} G_3(it_1, it_2) &= \int d_3^{1/2}\sigma^3_{-t_2/2}(a)\,\sigma^3_{t_1/2}(b)\,d_3^{1/2}K_3(\xi)\,d\omega_3' = \\ &\int d_3^{1/2}K_3(\xi)\,d_3^{1/2}\sigma^3_{-t_2/2}(a)\,\sigma^3_{t_1/2}(b)\,d\omega_3' = \\ &\int d_2^{1/2}\,\mathsf{E}\,(K_3(\xi))\,d_2^{1/2}\sigma^2_{-t_2/2}(a)\,\sigma^2_{t_1/2}(b)\,d\omega_2' = \\ &\int d_2^{1/2}K_2(\xi)\,d_2^{1/2}\sigma^2_{-t_2/2}(a)\,\sigma^2_{t_1/2}(b)\,d\omega_2' = \\ &G_2(it_1, it_2). \end{aligned}$$

Here 2.7 Lemma from 3 was used to establish the equality

$$\int d_3^{1/2}xd_3^{1/2}yd\omega_3' = \langle \pi_3(x)\Phi, J_3\pi_3(y^*)\Phi\rangle =$$

$$\langle \pi(\mathsf{E}(x))\Phi, J_2\pi_2(y^*)\Phi\rangle = \int d_2^{1/2}\,\mathsf{E}(x)\,d_2^{1/2}yd\omega_2'$$

for some $x\in M_3$, $y\in M_1$ and $\mathsf{E}(K_3(\xi))=K_2(\xi)$ was proved in [5] on p. 60.

The computation in the other cases is quite similar.

$$G_3(1+it_1, 1+it_2) = \int \sigma^3_{-t_2/2}(a)\, d_3 \sigma^3_{t_1/2}(b)\, K_3(\xi)\, d\omega'_3 =$$
$$\int d_3^{1/2} \sigma^3_{t_1/2}(b)\, K_3(\xi)\, \sigma^3_{-t_2/2}(a)\, d_3^{1/2}\, d\omega'_3 =$$
$$\int d_2^{1/2}\, \mathrm{E}\big(\sigma^2_{t_1/2}(b)\, K_3(\xi)\, \sigma^2_{-t_2/2}(a)\big)\, d_2^{1/2}\, d\omega'_2 =$$
$$\int d_2^{1/2} \sigma^2_{t_1/2}(b)\, K_2(\xi)\, \sigma^2_{-t_2/2}(a)\, d_2^{1/2}\, d\omega'_2 =$$
$$G_2(1+it_1, 1+it_2)$$

$$G_3(1+it_1, it_2) = \int d_3^{1/2} \sigma^3_{-t_2/2}(a)\, d_3^{1/2} \sigma^3_{t_1/2}(b)\, K_3(\xi)\, d\omega'_3 =$$
$$\int d_2^{1/2} \sigma^2_{-t_2/2}(a)\, d_2^{1/2} \sigma^2_{t_1/2}(b)\, K_2(\xi)\, d\omega'_2 =$$
$$G_2(1+it_1, it_2)$$

$$G_3(it_1, 1+it_2) = \int \sigma^3_{-t_2/2}(a)\, d_3^{1/2} \sigma^3_{t_1/2}(b)\, d_3^{1/2} K_3(\xi)\, d\omega'_3 =$$
$$\int \sigma^2_{-t_2/2}(a)\, d_2^{1/2} \sigma^2_{t_1/2}(b)\, d_2^{1/2}\, K_2(\xi)\, d\omega'_2 =$$
$$G_2(it_1, 1+it_2).$$

(vi)$\to$(v): Since the particular case $p=1$ and $s=\infty$ of (vi) is essentially (v).

COROLLARY. *Assume that there is a conditional expectation F of M_3 onto M_2. Then the conditions in the Theorem are equivalent to the following one:*

(vii) *For $t \in \mathbf{R}$ there is a unitary $u_t \in M_3 \cap M_1'$ such that*

$$u_t a u_t^* = \sigma^3_{-t} \sigma^2_t(a) \quad \textit{for all} \quad a \in M_2.$$

PROOF. Let $\psi = \varphi_2 \circ F$ be a state on M_3. Choosing $u_t = [D\varphi_3, D\psi]_t$ we have

$$\sigma^3_t(a) = u_t \sigma^\psi_t(a)\, u_t^* \quad (a \in M_3).$$

But $\sigma^\psi_t | M_2 = \sigma^2_t$ ([9], 10.5) and this gives

$$u_t a u_t^* = \sigma^3_{-t} \sigma^2_t(a) \qquad (a \in M_2).$$

Assuming (iv)$'$ we infer that $u \in M_1'$.

On the other hand, (vii) implies (iv)$'$ easily putting $a \in M_1$.

REFERENCES

[1] ACCARDI, L. and CECCHINI, C., Conditional expectations in von Neumann algebras and a theorem of Takesaki, *J. Funct. Anal.* **45** (1982), 245—273. *Zbl* **483**. 46043; *MR* **84j**: 46088.

[2] ACCARDI, L. and FRIGERIO, A., Markovian cocycles, *Proc. Roy. Irish Acad. Sect A* **83** (1983), 251—263. *MR* 85g: 46078.

[3] CECCHINI, C., Non-commutative integration for states on von Neumann algebras, *J. Operator Theory* **15** (1986), 217—237.

[4] CECCHINI, C., Non-commutative L^p-spaces and KMS functions, *Quantum Probability and Applications*, Lecture Notes in Math., Vol. 1136, Springer, 1985, 136—142.

[5] CECCHINI, C. and PETZ, D., Norm convergence of generalized martingales in L^p-spaces over von Neumann algebras, *Acta Sci. Math. (Szeged)* **48** (1985), 55—63. *MR* **87m**: 46125
[6] CONNES, A., On the spatial theory of von Neumann algebras, *J. Funct. Anal.* **35** (1980), 153—164. *MR* **81g**: 46083.
[7] HILSUM, M., Les espaces L^p d'une algèbre de von Neumann définies par la dérivée spatiale, *J. Funct. Anal.* **40** (1981), 151—169. *MR* **83c**: 46053.
[8] PETZ, D., A dual in von Neumann algebras with weights, *Quart. J. Math. Oxford* **35** (1984), 475—483. *MR* **86c**: 46080
[9] STRĂTILĂ, S., *Modular theory in operator algebras*, Abacus Press, Tunbridge Wells, 1981. *MR* **85g**: 46072.
[10] TAKESAKI, M., Conditional expectations in von Neumann algebras, *J. Functional Analysis* **9** (1972), 306—321. *MR* **46** # 2445.

(Received March 13, 1986)

ISTITUTO DI MATEMATICA
UNIVERSITÀ DI GENOVA
VIA L. B. ALBERTI 4
I—16132 GENOVA
ITALY

MTA MATEMATIKAI KUTATÓ INTÉZETE
POSTAFIÓK 127
H—1364 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 139—146

A NOTE ON THE CATEGORY OF S.M.F. SPACES AND RELATED CATEGORIES

V. S. KRISHNAN

Abstract

The spaces with a directed family of semi-metrics, are the objects of two categories $\mathcal{SMF}$ and $\mathcal{SMF}^*$; the difference is in the family of morphisms, which are "contractions" in $\mathcal{SMF}$ and near-contractions in $\mathcal{SMF}^*$. Correspondingly we have $\mathcal{SM}$ and $\mathcal{SM}^*$, the categories of semi-metric spaces with contractions or near-contractions for morphisms. The main results proved here are: Theorem 1 which says that there is a coretraction of the category $\mathcal{SU}$ of semi uniform spaces on $\mathcal{SMF}$, and an equivalence of the category $\mathcal{SSU}$, of strong semiuniform spaces with $\mathcal{SMF}^*$. The second theorem establishes equivalences: one of $\mathcal{SMF}$ with a subcategory of the procategory over $\mathcal{SM}$, and another of $\mathcal{SMF}^*$ with a subcategory of the procategory over $\mathcal{SM}^*$.

Introduction

The s.m.f. structures or semi-metric family structures were introduced in an earlier paper [2] in connection with semigroups. It was noted there that these s.m.f. spaces occupy a position in between the semimetric spaces and the semiuniform spaces. This paper establishes two main lines of connection (functorial ones): there are functors U, M between the categories of s.m.f. spaces and the s.u. spaces which give a contraction of $\mathcal{SU}$ in $\mathcal{SMF}$. But by enlarging the sets of morphisms in these categories, we get categories $\mathcal{SSU}$ and $\mathcal{SMF}^*$ which are actually equivalent. This is the first theorem. The next one relates the categories $\mathcal{SMF}$ and $\mathcal{SMF}^*$ to Procategories over the corresponding categories $\mathcal{SM}$ and $\mathcal{SM}^*$ (of semimetric spaces with contraction maps or near-contraction maps for their morphisms). Here we get again an equivalence of $\mathcal{SMF}$ or $\mathcal{SMF}^*$ with a certain subcategory $\mathcal{S}$ Pro $\mathcal{SM}$ or $\mathcal{S}$ Pro $\mathcal{SM}^*$ of the pro-categories. This is the second theorem.

1. The categories, $\mathcal{SM}$, $\mathcal{SM}^*$, $\mathcal{SMF}$, $\mathcal{SMF}^*$, $\mathcal{SU}$ and $\mathcal{SSU}$

A *semi-metric* d on a set X is a mapping of $X\times X$ in the real unit interval $I=[0, 1]$ such that: for all x of X, $d(x, x)=0$, and for all x, y, z from X, $d(x, y)+d(y, z)\geqq d(x, z)$. A *semi-metric space* (or *s.m. space*) (X, d) is a set X provided with a semi-metric d on it. Given two s.m. spaces (X, d) and (Y, d'), a map $f\colon X\to Y$ is called a (i) *contraction*, (ii) a *near-contraction* (of index r), of (X, d) in (Y, d'),

---

1980 *Mathematics Subject Classification*. Primary 54E99.

Key words and phrases. Semimetric spaces, semimetric family spaces, contractions, near-contractions, semiuniform spaces and strong semiuniform spaces, procategories over the categories of semimetric spaces.

if (i) for all x_1, x_2 from X, $d(x_1, x_2) \geqq d'(f(x_1), f(x_2))$, (ii) there is a positive real number r such that, for all x_1, x_2 from X, $d(x_1, x_2) \geqq rd'(f(x_1), (f(x_2))$. We have then two categories $\mathscr{SM}$ and $\mathscr{SM}^*$ both of them having the s.m. spaces for the objects; the morphisms for $\mathscr{SM}$ are the contractions, while those of $\mathscr{SM}^*$ are the near-contractions, between pairs of s.m. spaces.

When J is a down-directed (ordered) set and for each j of J there is assigned a semi-metric d_j on the set X such that, when $j \leqq j'$, the map I_X is a contraction of (X, d_j) in $(X, d_{j'})$, we call (d, J) a *semi-metric family structure* (or *s.m.f. structure*) for X, and also call (X, d, J) a *semi-metric family space* (or *s.m.f. space*). Given two such spaces (X, d, J) and (Y, d', K), a map $f\colon X \to Y$ is called a (i) *contraction*, (ii) a *near-contraction* of the first space in the second if (i) there is a map $f^*\colon K \to J$ (called a comap of f) such that, when $j = f^*(k)$, f is a contraction of (X, d_j) in (Y, d'_k), (ii) there is a map $f^*\colon K \to (J \times R^+)$ (called a comap of f) such that, when $f^*(k) = (j, r)$, for all x_1, x_2 of X, $d_j(x_1, x_2) \geqq r \cdot d'_k(f(x_1), f(x_2))$. That is, f is a near-contraction of (X, d_j) in (Y, d_k) of index r, when $f^*(k) = (j, r)$. We get again two categories $\mathscr{SMF}$ and $\mathscr{SMF}^*$, both having the same family of objects, the s.m.f. spaces; the morphisms for $\mathscr{SMF}$ are the contractions, while for $\mathscr{SMF}^*$ they are the near-contractions, between pairs of s.m.f. spaces.

When J is a down-directed set, a monotone map U of J in the power set $P(X \times X)$ of all subsets of $X \times X$, gives a semi-uniform structure (U, J) for X if: for each j of J, $U(j)$ contains the identity relation I_X, and also contains the relational product $U(j') \circ U(j')$ for some j' of J. We call (X, U, J) a *semi-uniform space* (or *s.u. space*) when (U, J) is a semi-uniform structure on X. Given two such spaces (X, U, J) and (Y, V, K) we call a map $f\colon X \to Y$ a *uniform map* of the first space in the second if there is a map $f^*\colon K \to J$, (called a comap of f) such that: for all k of K and for all x_1, x_2 from X, $(x_1, x_2) \in U(f^*(k))$ implies $(f(x_1), f(x_2)) \in V(k)$. We denote this also by writing $f[U(f^*(k))] \subseteq V(k)$, for all k of K. The category $\mathscr{SU}$ has semiuniform spaces for objects and uniform maps between pairs of them for morphisms. By a *strong semiuniform structure* (U, J, c) on the set X we mean a triple consisting of a down-directed set J, a monotone map U of J in $P(X \times X)$, and a monotone map c of J in J, such that for each j of J, $U(j)$ contains I_X and the relational fourth power $[U(cj)]^4$, where we set cj for $c(j)$, and similarly $c^2 j$ for $c(cj)$, etc. A strong semiuniform structure (U, J, c) on X leads to a *strong semiuniform space* (X, U, J, c). Given two such spaces (X, U, J, c) and (Y, V, K, c) a map $f\colon X \to Y$ is called a *strong uniform map* of the first in the second space if there is a map $f^*\colon K \to (J \times N^*)$ such that, when $f^*(k) = (j, m)$: for each n in N^*(=the set of integers$\geqq 0$), $f[U(c^{n+m} j)] \subseteq \subseteq V(c^n k)$. We have then a category $\mathscr{SSU}$ whose objects are the strong semiuniform spaces and whose morphisms are the strong uniform maps between pairs of these spaces.

2. The categories Pro $\mathscr{SM}$, Pro $\mathscr{SM}^*$, $\mathscr{S}$ Pro $\mathscr{SM}$ and $\mathscr{S}$ Pro $\mathscr{SM}^*$

We define the category Pro C over a category C essentially as in the book on Shape Theory [3, § 1] except, that we take the indexing sets to be down-directed rather than up-directed, which involves turning the order relation around in all the basic notions.

Given a down directed (indexing) set J, and for each j of J an object X_j of the category C, and for each pair j, j' from J with $j \leqq j'$ a morphism $p_{j,j'}$ of C from X_j to $X_{j'}$ such that (i) $p_{j,j} = I_{X_j}$ for each j, and (ii) $p_{j',j''} p_{j,j'} = p_{j,j''}$, whenever $j \leqq j' \leqq j''$ in J, we call the triple $(X_j, p_{j,j'}, J)$ a *directed family over* C. Given two such directed families $(X_j, p_{j,j'}, J)$ and $(Y_k, q_{k,k'}, K)$ a *morphism* of the first in the second is a pair (f_k, φ) consisting of a map φ of K in J and, for each k of K, a morphism f_k, in C, from $X_{\varphi(k)}$ in Y_k, such that: whenever $k \leqq k'$ in K, we have the following commutative relation, for some j of J which is $\leqq$ both $\varphi(k)$ and $\varphi(k')$ in J: $q_{k,k'} f_k p_{j,\varphi(k)} = f_{k'} p_{j,\varphi(k')}$. When (f_k', φ') is another such morphism of $(X_j, p_{j,j'}, J)$ in $(Y_k, q_{k,k'}, K)$, we call the two morphisms (f_k, φ) and (f_k', φ') equivalent if, for each k of K there is a j of J which is $\leqq$ both $\varphi(k)$ and $\varphi'(k)$, such that we have the commutativity relation: $f_k p_{j,\varphi(k)} = f_k' \cdot p_{j,\varphi'(k)}$. It is not hard to see that this is indeed an equivalence relation among morphisms between the pair of directed families; and further, if we have a second pair of equivalent morphisms (g_1, θ) and (g_1', θ') from $(Y_k, q_{k,k'}, K)$ to $(Z_1, r_{1,1'}, L)$, then the composite morphism $(g_1 \cdot f_{\theta(1)}, \varphi. \theta)$ and $(g_1' \cdot f_{\theta'(1)}', \varphi' \cdot \theta')$ both from $(X_j, p_{j,j'}, J)$ to $(Z, r_{1,1'}, L)$ are also equivalent. So we now define the category Pro C to consist of: the directed families over C for its objects, and equivalence classes of morphisms between pairs of such families for the morphisms. The composite of two equivalence classes can be taken as the equivalence class containing the composite of representative morphisms taken one from each of the equivalence classes. The full subcategory of Pro C whose objects $(X_j, p_{j,j'}, J)$ satisfy the extra condition: each of the morphisms $p_{j,j'}$ (for $j \leqq j'$ in J) is both an epi and mono-morphism, is denoted by $\mathscr{S}$ Pro C; these objects are called *special directed families* over C. The two cases we consider here are when C is either the category $\mathscr{SM}$ or the category $\mathscr{SM}^*$. For these cases, the special directed family $(X_j, p_{j,j'}, J)$ can be more simply described by saying that the maps $p_{j,j'}$ (for $j \leqq j'$ in J) are all one-one and onto, as set-maps.

3. Functorial connections of $\mathscr{SMF}$ with $\mathscr{SU}$ and of $\mathscr{SMF}^*$ with $\mathscr{SSU}$

THEOREM 1. (a) *There are functors* U: $\mathscr{SMF} \to \mathscr{SU}$ *and* M: $\mathscr{SU} \to \mathscr{SMF}$ *such that* M *is a left-adjoint of* U, *and* U · M *is naturally equivalent with* $I_{\mathscr{SU}}$.

(b) *There are functors* U*: $\mathscr{SMF}^* \to \mathscr{SSU}$ *and* M*: $\mathscr{SSU} \to \mathscr{SMF}^*$ *such that* M* *is a left-adjoint of* U*, *and* U*, M* *give an equivalence between the categories* $\mathscr{SMF}^*$ *and* $\mathscr{SSU}$.

PROOF. (a) For an object (X, d, J) of $\mathscr{SMF}$ we set $U(X, d, J) = (X, U(J \times N^*))$, where N^* is the set of integers $\geqq 0$ ordered by $\geqq$, and $U(j, n)$ is defined to be $[(x, y) \in X \times X: d_j(x, y) < 1/2^n]$. Clearly each $U(j, n)$ contains I_X and $U(j, n+1) \circ \circ U(j, n+1)$. For a morphism $f: (X, d, J) \to (Y, d', K)$ of $\mathscr{SMF}$ (with comap $f^*: K \to J$), we associate $U(f)$ = the same set-map f considered as a morphism in $\mathscr{SU}$ from $U(X, d, J)$ to $U(Y, d', K)$ with comap $f\#$ given by $f\#(k, n) = (f^*(k), n)$. Since $d_j(x, y) \geqq d_k'(f(x), f(y))$, when $j = f^*(k)$, this is a valid definition of a morphism from $\mathscr{SU}$. That this gives a functor U as asserted from $\mathscr{SMF}$ to $\mathscr{SU}$ is easily verified (U takes identity morphisms to identity morphisms, and composites to composites.)

For an object (X, U, J) of $\mathscr{SU}$ we set $M(X, U, J) = (X, d^*, J^*)$ where $[d_{j^*}^*:$ j^* in $J^*]$ is the family of all semimetrics on X such that: for each positive integer n, and j^* of J^* there is a j of J such that (x, y) in $U(j)$ implies that $d_{j^*}^*(x, y) < 1/2^n$; in other words, the semi-uniform structure on X determined by the semi-metric $d^*_{j^*}$ (for each j^* of J^*) is coarser than the semiuniform structure (U, J) on X. Again when f is a morphism in $\mathscr{SU}$ from (X, U, J) to (Y, V, K) we can set $\mathrm{M}(f)$ = the same set map f considered as a morphism in $\mathscr{SMF}$ from $\mathrm{M}(X, U, J)$ to $\mathrm{M}(Y, V, K)$. For if f^* is the comap for the original f from $\mathscr{SU}$, for a k^* of K^*, we can associate a $j^* = f\sharp(k^*)$ of J^* by setting: $d^*_{j^*}(x, y) = d^{*\prime}_{k^*}(f(x), f(y))$. For an integer n $(\geqq 0)$ we know there is a k in K such that $(f(x), f(y))$ in $V(k)$ implies that $d^{*\prime}_{k^*}(f(x), f(y)) < < 1/2^n$, and for this k, and $j = f^*(k)$, (x, y) in $U(j)$ implies that $(f(x), f(y))$ is in $V(k)$, hence $d^*_{j^*}(x, y) < 1/2^n$. Hence this j^* is in J^*, and we get the comap $f\sharp$ of f, to make f a morphism in $\mathscr{SMF}$. The rest of the verification that this M is indeed a functor from $\mathscr{SU}$ to $\mathscr{SMF}$ is easy.

For an object (X, U, J) of $\mathscr{SU}$ it is clear, from the definitions, that I_X is a uniform map of (X, U, J) in (on) $\mathrm{U} \cdot \mathrm{M}(X, U, J) = U[(X, d^*, J^*)] = (X, U^*, (J^* \times N^*))$. The other way, I_x is still a uniform map of $(X, U^*, (J^* \times N^*))$ in (X, U, J): first, to each j of J we associate a $c(j)$ or cj of J such that $U(j)$ contains $[U(cj)]^4$, and in terms of the sequence $[U(j), U(cj), U(c^2j), \ldots]$ we can define on X a semi-metric d^*_j such that:

$$(x, y) \in U(c^{n+1}j) \Rightarrow d^*_j(x, y) \leqq 1/2^{n+1} < 1/2^n \Rightarrow (x, y) \in U(c^n j).$$

This process we work out in the proof of the next result (b). Then $U(j)$ contains $U^*(j, 0)$, associated with this d^*_j, and this d^*_j is one of the $d^*_{j^*}(j^*$ in $J^*)$ in the definition of $\mathrm{M}(X, U, J) = (X, D^*, j^*)$. These show that $\mathrm{U} \cdot \mathrm{M}$ is naturally equivalent to $I_{\mathscr{SU}}$.

For an object (X, d, J) of $\mathscr{SMF}$, if $U(X, d, J) = (X, U(J \times N^*))$, each d_j, j in J, gives a semiuniform structure on X coarser than $(U, (J \times N^*))$, so can be considered as one of the $d^*_{j^*}, j^*$ in J^* from the $(X, d^*, J^*) = \mathrm{M}(X, U, (J \times N^*)) = = \mathrm{M} \cdot \mathrm{U} \cdot (X, d, J)$. Thus I_X is a morphism in $\mathscr{SMF}$ from $\mathrm{M} \cdot \mathrm{U}(X, d, J)$. This means that M is a left-adjoint of U.

This completes the proof of (a).

(b) For an object (X, d, J) of $\mathscr{SMF}^*$ we set $\mathrm{U}^*(X, d, J) = (X, U, (J \times N^*), c)$, where $U(j, n) = [(x, y) \in X \times X : d_j(x, y) < 1/2^n]$, and $c(j, n) = (j, n+2)$. This ensures that $U(j, n)$ contains $[U(c(j, n))]^4$. For a morphism f from (X, d, J) to (Y, d', K) in $\mathscr{SMF}^*$ with comap f^*, we set $\mathrm{U}^*(f)$ = the same set-map f of X in Y considered as a morphism in $\mathscr{SSU}$ from $\mathrm{U}^*(X, d, J)$ to $\mathrm{U}^*(Y, d', K)$. For if $f^*(k) = = (j, r)$, for a k of K, we can set $f\sharp\,(k, n) = [(j, n), m]$, with m as the smallest integer such that $2^m \geqq 1/r$; this $f\sharp$ would be a comap for f (in $\mathscr{SSU}$). For given any positive integer n, $(x, y) \in U(c^{m+n}j)$ implies that $d_j(x, y) < 1/2^{n+m} \leqq r \cdot 1/2^n$; from $d_j(x, y) \geqq r d_k(f(x), f(y))$ we deduce that $d_k(f(x), f(y)) \leqq (1/r) d_j(x, y) < 1/2^n$, so $(f(x), f(y)) \in V(c^n k)$. This proves that f is indeed a morphism in $\mathscr{SU}$ (with comap $f\sharp$).

We next define M^* from $\mathscr{SSU}$ to $\mathscr{SMF}^*$. For an object (X, U, J, c) from $\mathscr{SSU}$, we associate an object $\mathrm{M}^*(X, U, J, c) = (X, d, J)$ by defining, for each j of J a semimetric d_j on X as follows: first we define an auxiliary function g from $X \times X$ to I by: $g(x, y) = 0$ if (x, y) belongs to $U(c^n j)$ for all integers $n \geqq 0$, $g(x, y) = 1$ if (x, y) is not even in $U(c^0 j) = U(j)$, and $g(x, y) = 1/2^n$ if (x, y) belongs to $U(c^n j)$ but not to $U(c^{n+1} j)$. Since the sets $U(c^n j)$ form a decreasing sequence as n increases, this provides a $g(x, y)$ defined for all (x, y) from $X \times X$.

Then we set $d_j(x, y) = \inf[\Sigma g(x_i, x_{i+1}), \ i = 1, \ldots, n$: these sums being taken over all sequences $x=x_1, \ldots, x_{n+1}=y$ of points connecting x with y]. It is not hard to show that: for all x, y from X, $(x, y)\in U(c^n, y)$ iff $g(x, y)\leqq 1/2^n$, and that $1/2g(x, y)\leqq d_j(x, y)\leqq g(x, y)$. [See for proofs [1], page 28, proof of theorem 1.1.5.] It follows from these that $(x, y)\in U(c^{n+1}j)\Rightarrow d_j(x, y)<1/2^n\Rightarrow(x, y)\in U(c^n j)$ ... (A). These semimetrics do give an object (X, d, J) of $\mathscr{SMF}^*$, for it can be seen that: $j\leqq j'$ in J implies that $d_j(x, y)\geqq d_{j'}(x, y)$, for all x, y from X, since $g(x, y)\geqq g'(x, y)$ if g and g' are the auxiliary functions from $X\times X$ to I defined in terms of the $[U(c^n j)]$ and the $[U(c^n j')]$, resply. For a morphism f in $\mathscr{SSU}$ from (X, U, J, c) and (Y, V, K, c) with comap f^*, we associate $\mathrm{M}^*(f)=$the same set map f of X in Y considered as a morphism in $\mathscr{SMF}^*$ from $\mathrm{M}^*(X, U, J, c)$ to $\mathrm{M}^*(Y, V, K, c)$ with a comap $f\#$: where $f\#(k)$ is defined as $(j, 1/2^m)$, when $f^*(k)=(j, m)$. What has to be verified is that, when $f^*(k)=(j, m)$, $d_j(x, y)\geqq 1/2^m\cdot d_k(f(x), f(y))$, for each pair (x, y) from $X\times X$. This being trivially true if either $d_j(x, y)=0$ or $d_j(x, y)\geqq 1/2^m$, we take the case when $0\neq d_j(x, y)<1/2^m$. Then, from the definition of d_j, there is a finite sequence $x=x_1, \ldots, x_{n+1}=y$ from x to y such that $d_j(x, y)\leqq \Sigma g(x_i, x_{i+1})<<1/2^n$. So each of the $g(x_i, x_{i+1})$ is different from 0 and less than $1/2^m$. Let $g(x_i, x_{i+1})=1/2^{n_i}$. Each $n_i>m$; and as $f^*(k)=(j, m)$, $(x_i, x_{i+1})\in U(c^{n_i}j)$ implies that $(f(x_i), f(x_{i+1}))\in V(c^{n_i-m}k)$, which gives $d_k[f(x_i), f(x_{i+1})]<1/2^{n_i-m}$; hence $d_k(f(x), f(y))<2^m\cdot\Sigma 1/2^{n_i}=2^m\cdot\Sigma g(x_i, x_{i+1})$. It follows that $1/2^m\cdot d_k(f(x), f(y))$, being less than or equal to each of these sums $\Sigma g(x_i, x_{i+1})$ which have $d_j(x, y)$ for infimum, must be also less than or equal to $d_j(x, y)$.

The rest of the verification that this M^* is indeed a functor from $\mathscr{SSU}$ to $\mathscr{SMF}^*$ is pretty routine.

To prove the equivalence of the categories $\mathscr{SSU}$ and $\mathscr{SMF}^*$ under the functors U^*, M^* we prove that $\mathrm{U}^*\mathrm{M}^*$ is naturally equivalent to $I_{\mathscr{SSU}}$ and $\mathrm{M}^*\mathrm{U}^*$ is naturally equivalent to $I_{\mathscr{SMF}^*}$. Starting from an object (X, U, J, c) of $\mathscr{SSU}$, let M^* take it to the object (X, d, J) of $\mathscr{SMF}^*$ and let $\mathrm{U}^*(X, d, J)=(X, U^*, (J\times N^*), c)$. We show that (I_X, f^*) is a morphism of (X, U, J, c) on $(X, U^*, (J\times N^*), c)$ and $(I_X, f\#)$ a morphism of $(X, U^*, (J\times N^*), c)$ on (X, U, J, c), where f^* assigns $(j, m+1)$ to the element (j, m) from $(J\times N^*)$, and $f\#$ assigns $((j, 0), 0)$ to the element j from J. For the first, note that $(x, y)\in U(c^{n+m+1}j)$ implies $d_j(x, y)<1/2^{n+m}$ which implies $(x, y)\in U^*(j, n+m)=U^*[c^n(j, m)]$. For the second note that $(x, y)\in U^*(c^{n+0}(j, 0))==U^*(j, n)$ implies that $d_j(x, y)<1/2^n$ which implies that $(x, y)\in U(c^n j)$. These prove that there is a natural equivalence from $I_{\mathscr{SSU}}$ to $\mathrm{U}^*\mathrm{M}^*$ which assigns to an object (X, U, J, c) of $\mathscr{SSU}$ the morphism (I_X, f^*) as defined above. The other way, starting from an object (X, d, J) of $\mathscr{SMF}^*$, let U^* take it to $(X, U, (J\times N^*), c)$ and M^* take this $(X, U^*, (J\times N^*), c)$ to $(X, d^*, (J\times N^*))$. We establish a natural transformation which is an equivalence from $\mathrm{M}^*\cdot\mathrm{U}^*$ to $\mathrm{I}_{\mathscr{SMF}}$; this associates to the object (X, d, J) of $\mathscr{SMF}^*$ the morphism (I_X, g^*) from $(X, d^*, (J\times N^*))$ to (X, d, J) with a reverse morphism $(I_X, g\#)$ from (X, d, J) to $(X, d^*, (J\times N^*))$: where $g^*(j)=[(j, 0), 1]$ and $g\#(j, n)=[j, 1/2^n]$. To prove that (I_X, g^*) is a morphism as claimed, note that $d^*_{(j,0)}(x, y)=d_j(x, y)$, when $d_j(x, y)=1$; since, in this case $(x, y)\notin U(j)=U(c^0(j, 0))$. If $d_j(x, y)<1$, there are sums $\Sigma g(x_i, x_{i+1})$ for sequences of points from x to y which are <1 and $\geqq d_j(x, y)$. For any such sum, if $g(x_i, x_{i+1})=1/2^{n_i}$, $d_j(x_i, x_{i+1})\leqq g(x_i, x_{i+1})$, for each of these pairs (x_i, x_{i+1}), so that $d_j(x, y)\leqq\Sigma d_j(x_i, x_{i+1})\leqq\Sigma g(x_i, x_{i+1})$. Hence $d_j(x, y)\leqq$the infimum of such sums which is $d^*_{(j,0)}(x, y)$. To prove that $(I_X, g\#)$ is a morphism as claimed, we have to show that, for all x, y from X, $d_j(x, y) \geqq 1/2^{m+1} \cdot d^*_{(j,m)}(x, y)$. The proof is pretty obvious if $d_j(x, y)=0$, or if $d_j(x, y) \geqq 1/2^{m+1}$. So we assume that $d_j(x, y) \neq 0$ and is less than $1/2^{m+1}$. Then $d_j(x, y)$ lies in an interval of t of the form $[1/2^{r+1}, 1/2^r)$ with $r \geqq m+1$. So (x, y) is in $U(c^r j)$ but not in $U(c^{r+1}j)$, that is in $U[c^{r-m}(c^m j)]$ but not in $U[c^{r-m+1}(c^m j)]$. This means, that, for the auxiliary function g defining $d^*_{(j,m)}$, $g(x, y)=1/2^{r-m}$; hence, $d^*_{(j,m)}(x, y) \leqq g(x, y) = 1/2^{r-m} = (1/2^{r+1}) \cdot 2^{m+1} \leqq \leqq d_j(x, y) 2^{m+1}$ since, by assumption, $d_j(x, y) \geqq 1/2^{r+1}$. This proves that M*, U* give indeed an equivalence for the categories $\mathcal{SSU}$ and $\mathcal{SMF}^*$.

4. Functional connections of $\mathcal{SMF}/\mathcal{SMF}^*$ with the categories Pro $\mathcal{SM}$, $\mathcal{S}$ Pro $\mathcal{SM}$/Pro $\mathcal{SM}^*$, $\mathcal{S}$ Pro $\mathcal{SM}^*$

THEOREM 2. (a) *There are functors* $\mathrm{P}: \mathcal{SMF} \to \mathcal{S}$ Pro $\mathcal{SM}$, $\mathrm{J}_p: \mathcal{S}$ Pro $\mathcal{SM} \to$ Pro $\mathcal{SM}$, *and* F: Pro $\mathcal{SM} \to \mathcal{SMF}$, *such that: there is a natural transformation from* $\mathrm{J}_p\mathrm{PF}$ *to* $\mathrm{I}_{\mathrm{Pro}\,\mathcal{SM}}$ *and natural equivalences, one from* $\mathrm{I}_{\mathcal{SMF}}$ *to* $\mathrm{FJ}_p\mathrm{P}$ *and one from* $\mathrm{I}_{\mathcal{S}\,\mathrm{Pro}\,\mathcal{SM}}$ *to* PFJ_p. *Thus* $\mathcal{SMF}$ *and* $\mathcal{S}$ Pro $\mathcal{SM}$ *are equivalent categories, and there are coretractions of these to* Pro $\mathcal{SM}$. (b) *There are functors* $\mathrm{P}^*: \mathcal{SMF}^* \to \mathcal{S}$ Pro $\mathcal{SM}^*$, $\mathrm{J}^*_p: \mathcal{S}$ Pro $\mathcal{SM}^* \to$ Pro $\mathcal{SM}^*$, *and* F^*: Pro $\mathcal{SM}^* \to \mathcal{SMF}^*$, *such that: there is a natural transformation from* $\mathrm{J}^*_p\,\mathrm{P}^*\mathrm{F}^*$ *to* $\mathrm{I}_{\mathrm{Pro}\,\mathcal{SM}^*}$ *and natural equivalences, one from* $\mathrm{I}_{\mathcal{SMF}^*}$ *to* $\mathrm{F}^*\mathrm{J}^*_p\mathrm{P}^*$ *and one from* $\mathrm{I}_{\mathcal{S}\,\mathrm{Pro}\,\mathcal{SM}^*}$ *to* $\mathrm{P}^*\mathrm{F}^*\mathrm{J}^*_p$. *Thus* $\mathcal{SMF}^*$ *and* $\mathcal{S}$ Pro $\mathcal{SM}^*$ *are equivalent categories, and there are coretractions of these on* Pro $\mathcal{SM}^*$.

PROOF. We define the pairs of functors P, P* first. Given an object (X, d, J) of $\mathcal{SMF}$/of $\mathcal{SMF}^*$ we set $\mathrm{P}(X, d, J)/\mathrm{P}^*(X, d, J)$ to be the object $[(X, d_j), I_X, J]$ of $\mathcal{S}$ Pro $\mathcal{SM}$/of $\mathcal{S}$ Pro $\mathcal{SM}^*$. This is a valid definition, since, for $j \leqq j'$ in J, I_X is a one-one onto contraction/one-one onto near contraction, of (X, d_j) on $(X, d_{j'})$. For a morphism f from one object (X, d, J) to another (Y, d', K), that is a contraction/a near-contraction, with comap f^*, we define $\mathrm{P}(f)/\mathrm{P}^*(f)$ to be the equivalence class containing the morphism (f_k, φ) in $\mathcal{S}$ Pro $\mathcal{SM}$/in $\mathcal{S}$ Pro $\mathcal{SM}^*$ from $\mathrm{P}(X, d, J)$ to $\mathrm{P}(Y, d', K)$/from $\mathrm{P}^*(X, d, J)$ to $\mathrm{P}^*(Y, d', K)$; with $\varphi(k)=j$, and $f_k=f$ for all k of K when $f^*(k)=j$ when $f^*(k)=(j, r)$. This is again a valid definition since f is a contraction/a near contraction of index r, from (X, d_j) to (Y, d'_k) as defined. The rest of the verification that P, P* are functors is quite routine.

We next define the pair of functors F, F*. Let $[(X_j, d_j), p_{j,j'}, J]$ be an object of Pro $\mathcal{SM}$/Pro $\mathcal{SM}^*$; thus the down directed set J indexes family of s.m. spaces (X_j, d_j) and there is a contraction/near contraction (of index r), of (X_j, d_j) in $(X_{j'}, d_{j'})$ when $j \leqq j'$ in J; these being subject to the two conditions. We first associate a space to the family of s.m. spaces $[(X_j, d_j): j \text{ in } J]$; $\mathfrak{X}$ is the set of all elements ξ from the cartesian product set $P(X_j: j \text{ in } J)$ for which it is true that: whenever $j \leqq j'$ in J, $p_{j,j'}\,\xi(j)=\xi(j')$, where $\xi(j)$ denotes the jth component of ξ in (X_j). And we define, for each j of J, a semi-metric d^*_j on this set by setting $d^*_j(\xi_1, \xi_2) = d_j[\xi_1(j), \xi_2(j)]$, this last being in the s.m. space (X_j, d_j). Then $(\mathfrak{X}, d^*, J)$ becomes an s.m.f. space which we now define as $\mathrm{F}([(X_j, d_j), p_{j,j'}, J])$/as $\mathrm{F}^*([(X_j, d_j), p_{j,j'}, J])$ Let now $[(Y_k, d'_k), q_{k,k'}, K]$ be another object of Pro $\mathcal{SM}$/of Pro $\mathcal{SM}^*$, with F/F* taking it to $(\mathfrak{Y}, d'^*, K)$. Let also (f_k, φ) be a morphism from the directed family $[(X_j, d_j), p_{j,j'}, J]$ to the directed family $[(Y_k, d'_k), q_{k,k'}, K]$. We then define an associated map f of the set $\mathfrak{X}$ in the set $\mathfrak{Y}$; we set $f(\xi)=\eta$, where, for each k of K, $\eta(k)= =f_k[\xi(\varphi(k))]$. This f is then a morphism in $\mathscr{SMF}/\mathscr{SMF}^*$ from $(\mathfrak{X}, d^*, J)$ to $(\mathfrak{Y}, d'^*, K)$, with a comap f^* which takes a k of K to j/to (j, r), when $\varphi(k)=j$/ and f_k is a near contraction of index r. This follows from our definitions. Further, any other morphism (f_j, φ') equivalent to (f_k, φ) leads to the same morphism f: for, given k of K, we know that there is a j of J which is $\leqq$ both $\varphi(k)$ and $\varphi'(k)$ with $f_k p_{j,\varphi(k)}=f_k' p_{j,\varphi'(k)}$; and then $f(\xi)=\eta$, with $\eta(k)=f_k[\xi(\varphi(k))]=(f_k \cdot p_{j,\varphi(k)})\cdot[\xi(j)]$; while the f' defined by (f_k', φ') is given by $f'(\xi)=\eta'$, with $\eta'(k)=f_k'[\xi(\varphi'(k))]= =(f_k' p_{j,\varphi'(k)})[\xi(j)]$. Thus $f(\xi)=f'(\xi)$, for all ξ of X. Thus we can define, for the equivalence class of morphisms containing (f_k, φ) the associate under F/under F*, to be this morphism f in $\mathscr{SMF}/\mathscr{SMF}^*$. This completes the definition of the functors F/F*.

The functors $\mathrm{J}_p/\mathrm{J}_p^*$ are the obvious inclusion functors from the subcategories $\mathscr{S}$ Pro $\mathscr{SM}$ or $\mathscr{S}$ Pro $\mathscr{SM}^*$ to the full ones of Pro $\mathscr{SM}$ or Pro $\mathscr{SM}^*$. We tackle the proofs of parts (a) and (b) together; first, to get a natural transformation from $\mathrm{J}_p\mathrm{PF}$ to $\mathrm{I}_{\mathrm{Pro}\,\mathscr{SM}}$ and one from $\mathrm{J}_p^*\mathrm{P}^*\mathrm{F}^*$ to $\mathrm{I}_{\mathrm{Pro}\,\mathscr{SM}^*}$: given an object $[(X_j, d_j), p_{j,j'}, J]$ from Pro $\mathscr{SM}$ or from Pro $\mathscr{SM}^*$, this natural transformation would associate to it the morphism $[(f_j, \mathrm{I}_j)]$ in Pro $\mathscr{SM}$/Pro $\mathscr{SM}^*$ from $[(\mathfrak{X}, d_j^*), I_x, J]$ to $[(X_j, d_j), p_{j,j'}, J]$, where F/F* takes $[(X_j, d_j), p_{j,j'}, J]$ to $(\mathfrak{X}, d^*, J)$ as described a little earlier; where, for any j of J, we set $f_j(\eta)=\eta(j)$, for each element j of J. It is not hard to verify that: this f_j gives an isometry of the space $(\mathfrak{X}, d_j^*)$ with a subspace of (X_j, d_j) for each j and the other properties that ensure that (f_j, I_j) is indeed a morphism of $[(\mathfrak{X}, d_j^*), I_X, J]$ to $[(X_j, d_j), p_{j,j'}, J]$, and then we pass on to the equivalence class $[(f_j, I_j)]$ containing this pair, which is the typical morphism of Pro $\mathscr{SM}$/Pro $\mathscr{SM}^*$.

Next we show that there is a natural equivalence from $\mathrm{I}_{\mathscr{SMF}}/\mathrm{I}_{\mathscr{SMF}^*}$ to $\mathrm{FJ}_p\mathrm{P}/\mathrm{F}^*J_p^*\mathrm{P}^*$. To an object (X, d, J) of $\mathscr{SMF}/\mathscr{SMF}^*$ this would associate an isomorphism (f, f^*) from (X, d, J) to $(\mathfrak{X}, d^*, J)$ where $\mathfrak{X}$ is defined from the directed family $[(X_j, d_j), I_X, J]=$the image under P/P* of (X, d, J). It is clear that now we have the typical element $x\#$ of $\mathfrak{X}$ is an element of the power set X^J with $x\#(j)=x$ for each j of J; thus the map f is what takes an x of X to such an $x\#$ of $\mathfrak{X}$ with $x\#(j)=x$ for all j. This being obviously a one-one onto map, it needs only be noted that under this $d_j(x, y)=d_j^*(x\#, y\#)$, for any x, y from X. So f^* takes j of J to j in J/to (j, l) in $(J\times R^+)$, and (f, f^*) is an isomorphism in $\mathscr{SMF}/\mathscr{SMF}^*$.

Third we tackle the natural equivalence from $\mathrm{I}_{\mathscr{S}\,\mathrm{Pro}\,\mathscr{SM}}/\mathrm{I}_{\mathscr{S}\,\mathrm{Pro}\,\mathscr{SM}^*}$ to $\mathrm{PFJ}_p/\mathrm{P}^*\mathrm{F}^*\mathrm{J}_p^*$. Given an object $[(X_j, d_j), p_{j,j'}, J]$ of $\mathscr{S}$ Pro $\mathscr{SM}/\mathscr{S}$ Pro $\mathscr{SM}^*$, we have to get an associated isomorphism (f_j, I_J) of this object on $[(\mathfrak{X}, d_j^*), I_X, J]$, which is the P/P*-associate of the $\mathrm{FJ}_p/\mathrm{F}^*\mathrm{J}_p^*$-associate (X, d^*, J) of the original object $[(X_j, d_j), p_{j,j'}, J]$. Since the $p_{j,j'}$, are all one-one and onto, we can, for each j of J, get a bijection f_j of X_j on $\mathfrak{X}$, with inverse $f_j^{-1}=$the jth projection. For an x_j of X_j, and for any j_0 of J, we need only set $f_j(x_j)=\xi$, with $\xi(j_0)=p_{j',j_0}\cdot(p_{j',j}^{-1})(x_j)$, for a $j'\leqq$both j and j_0. That this definition of ξ is independent of the choice of such a j' can be shown, and this f_j is indeed a bijective map of X_j on X. Also this f_j gives an isometry of (X_j, d_j) with $(\mathfrak{X}, d_j^*)$. This completes the proof of the natural equivalence here. Thus the theorem is completely proved.

REFERENCES

[1] KRISHNAN, V. S., *An introduction to category theory,* North-Holland Publishing Co., New York—Amsterdam, 1981. *MR* **82f**: 18001.
[2] KRISHNAN, V. S., Semigroups with s.m.f. structures, *Semigroup Forum* **28** (1984), 13—27. *MR* **85h**: 20085.
[3] MARDEŠIĆ, S. and SEGAL, J., *Shape theory. The inverse system approach,* North-Holland Mathematical Library, Vol. 26, North-Holland, Publishing Co., Amsterdam, 1982. *MR* **84b**: 55020.

(Received April 15, 1986)

DEPARTMENT OF MATHEMATICS
TEMPLE UNIVERSITY
PHILADELPHIA, PA 19122
U.S.A.

Current address:

12818, S. E. 41ST LANE
NO. 13104
BELLEVUE, WA 98006
U.S.A.

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 147—177

PREPROXIMITIES AND INTERNAL CHARACTERIZATIONS OF COMPLETE REGULARITY

J. DEÁK

This paper has a double aim: first, to give a new internal characterization of completely regular spaces, and secondly, to generalize the notion of a preproximity; the latter will not only serve as a tool in proving the first, but it seems also to be of some interest in its own right.

§ 1 deals with characterizations of complete regularity in terms of the existence of a subbase satisfying certain conditions. The most general one of such characterizations in the literature (i.e. the one imposing the weakest condition on the subbase) is due to Hamburger ([27], [22]). Theorem 1.9 (proved in 5.33) generalizes his result; examples are given to show that our theorem is not only seemingly more general. Example 1.14 refutes a plausible generalization of Theorem 1.9.

§ 2 gives the definition of a preproximity in the sense of Hamburger [23], and suggests a modification in the terminology: originally, a preproximity is a pair consisting of a family of sets and a binary relation on this family; we call, however, the relation itself a preproximity.

§ 3 and § 4 contain a great number of definitions and lemmas.

§ 5 is the core of the paper. Definition 5.4 generalizes the notion of a preproximity. (Generalized preproximities are not required to be symmetrical, so they can be brought into connexion with quasi-proximities and bitopologies instead of proximities and topologies; but this is not the main point.) According to Theorem 5.10, each generalized preproximity induces a quasi-proximity in a natural way (and a direct construction of the totally bounded quasi-uniformity compatible with this quasi-proximity is given); conversely, each quasi-proximity can be induced by preproximities. Theorems 5.15 and 5.26 show that no further generalization of the notion of a preproximity is possible without losing its basic properties.

§ 6 deals with the category of the generalized preproximities. (Continuity is defined in the most obvious way.)

§ 7 indicates how to introduce preproximities for certain proximity-like relations satisfying axioms weaker than those of a (quasi-)proximity.

TERMINOLOGY. A *topology* will be regarded as the collection of the closed sets. A *(sub)base* for a topology means a (sub)base for the closed sets. $\mathrm{cl}_{\mathcal{T}}$ denotes the

---

1980 *Mathematics Subject Classification*. Primary 54D15, 54E05; Secondary 54E55, 54E99, 54B30, 54G20.

Key words and phrases. Completely regular, closed subbase, internal characterization, preproximity, (quasi-)proximity, quasi-uniformity, bitopology, topological category, bi(co)reflective subcategory.

closure in the topology $\mathcal{T}$. A *completely regular* topology is not required to be T_0. A *bitopology* is an ordered pair of topologies on the same set. For relations r and s between subsets of the same fundamental set, r is *finer* than s (or s is *coarser* than r) if $s\subset r$. (The symbol $\subset$ means not necessarily proper containing.) The definitions of proximity-like relations will always be based on the notion "far", see 4.0 for details. $A\,\bar{r}\,B$ means that $A\,r\,B$ does not hold.

§ 1. Subbase characterizations of complete regularity

Let us start from the following internal characterization of completely regular spaces:

1.0 Theorem (Frink [20]). *A T_1-space $(X, \mathcal{T})$ is completely regular iff there is a base $\mathcal{S}$ for $\mathcal{T}$ such that $\mathcal{S}$ is closed for finite unions and finite intersections, and*

(i) *if $A, B\in\mathcal{S}$, $A\cap B=\emptyset$ then there are $A', B'\in\mathcal{S}$ such that $A'\cup B'=X$ and $A\cap B'=\emptyset=A'\cap B$;*

(ii) *if $B\in\mathcal{S}$ and $x\in X\setminus B$ then there is an $A'\in\mathcal{S}$ with $x\in A'$, $A'\cap B=\emptyset$.*

The necessity of the condition is obvious, even without T_1 (let $\mathcal{S}$ be the family of all the zero sets), so we shall only concern ourselves with the generalizations of the sufficiency.

The next result is due to Hamburger [27] (a slightly weaker version, with (ii) instead of (iv), can be found in [22], [23], [24], [25]).

1.1 Theorem. *A topological space $(X, \mathcal{T})$ is completely regular if there is a subbase $\mathcal{S}$ for $\mathcal{T}$ such that*

(iii) *if $A, B\in\mathcal{S}$, $A\cap B=\emptyset$ then there are finite collections $\mathcal{A}, \mathcal{B}\subset\mathcal{S}$ with $A\subset\bigcup\mathcal{A}$, $B\subset\bigcup\mathcal{B}$, and for each pair $A'\in\mathcal{A}$, $B'\in\mathcal{B}$, there is a finite covering $\mathcal{C}\subset\mathcal{S}$ of X such that either $A'\cap C=\emptyset$ or $B'\cap C=\emptyset$ whenever $C\in\mathcal{C}$;*

(iv) *if $B\in\mathcal{S}$ and $x\in X\setminus B$ then there is a finite collection $\mathcal{B}\subset\mathcal{S}$ with $B\subset\bigcup\mathcal{B}$, and for each $B'\in\mathcal{B}$, there is an $A'\in\mathcal{S}$ such that $x\in A'$ and $A'\cap B'=\emptyset$.*

Results lying between Theorems 1.0 and 1.1 were given in [47], [6], [32], [1].

1.2 Example. Replacing (iv) by (ii) would weaken Theorem 1.1: let $|X|=4$, $|\mathcal{S}|=5$, and $|S|=2$ for each $S\in\mathcal{S}$.

1.3 Developing the idea of (iii), let us consider the following sequence of conditions, where $(\mathrm{iii})_1$ means (iii):

$(\mathrm{iii})_n$ if $A, B\in\mathcal{S}$, $A\cap B=\emptyset$ then there are finite collections $\mathcal{A}, \mathcal{B}\subset\mathcal{S}$ with $A\subset\bigcup\mathcal{A}$, $B\subset\bigcup\mathcal{B}$, $\bigcup\mathcal{A}\cap\bigcup\mathcal{B}=\emptyset$ such that the conclusion of $(\mathrm{iii})_{n-1}$ holds for each pair $A'\in\mathcal{A}$, $B'\in\mathcal{B}$ [where the sets A' and B' in $(\mathrm{iii})_n$ play the role of A and B in $(\mathrm{iii})_{n-1}$].

We shall later see that

1.4 Theorem. *A topological space $(X, \mathcal{T})$ is completely regular if there is a subbase $\mathcal{S}$ for $\mathcal{T}$ satisfying $(\mathrm{iii})_n$ and (iv), where n may depend on the pair (A, B).*

1.5 Example. Theorem 1.4 is strictly more general than Theorem 1.1. Let $X=[0, 1]$,

$$\mathcal{S} = \{[0, x], [x, 1], \{x\}: x\in X\}\cup\{\{0, 1/2, 1\}, [1/5, 2/5]\cup[3/5, 4/5]\}.$$

$\mathcal{S}$ satisfies $(\mathrm{iii})_2$ and (ii), but not (iii) (consider the last two sets in the definition of $\mathcal{S}$).

A yet stronger result:

1.6 Theorem. *A topological space $(X, \mathcal{T})$ is completely regular if there is a subbase $\mathcal{S}$ for $\mathcal{T}$ satisfying* (iv) *and*

(iii)* *if $A, B\in\mathcal{S}$, $A\cap B=\emptyset$ then there are finite collections $\mathcal{A}$ and $\mathcal{B}$ with $A=\bigcup\mathcal{A}$, $B=\bigcup\mathcal{B}$, and for each pair $A'\in\mathcal{A}$, $B'\in\mathcal{B}$, there are $A'', B''\in\mathcal{S}$ and a finite covering $\mathcal{C}\subset\mathcal{S}$ of X such that $A'\subset A''$, $B'\subset B''$, and either $A''\cap C=\emptyset$ or $B''\cap C=\emptyset$ whenever $C\in\mathcal{C}$.*

1.7 Remarks. a) It does not change (iii)* if we replace $A=\bigcup\mathcal{A}, B=\bigcup\mathcal{B}$ by $A\subset\bigcup\mathcal{A}$, $B\subset\bigcup\mathcal{B}$.

b) $(\mathrm{iii})_n\Rightarrow(\mathrm{iii})^*$ (proof by induction).

1.8 Example. Theorem 1.6 is strictly stronger than Theorem 1.4. Let

$$X = \{x_i, y_i, z_i: i = 1, 2, 3\},$$

$$\mathcal{S} = \{\{x_1, x_2, x_3\}, \{y_1, y_2, y_3\}, \{x_1, z_1\}, \{y_1, z_1\}\}\cup\{\{x_i, y_j, z_k\}: i, j, k = 2, 3\}.$$

$\mathcal{S}$ satisfies the conditions of 1.6 (in (iii)*, let $\mathcal{A}$ and $\mathcal{B}$ consist of singletons, and check that two distinct points are always contained by disjoint sets $S_1, S_2\in\mathcal{S}$ such that each $x\in X\setminus(S_1\cup S_2)$ is in an $S(x)\in\mathcal{S}$ disjoint either from S_1 or from S_2). 1.4, however, cannot be applied to any subsystem of $\mathcal{S}$. Indeed, if $\mathcal{S}'\subset\mathcal{S}$ is a subbase for the same topology then $A, B\in\mathcal{S}'$ where $A=\{x_1, x_2, x_3\}$, $B=\{y_1, y_2, y_3\}$; now assume that $(\mathrm{iii})_n$ holds for A and B; then $A\subset\bigcup\mathcal{A}$, $B\subset\bigcup\mathcal{B}$, $\bigcup A\cap\bigcup\mathcal{B}$ imply $A\in\mathcal{A}$ and $B\in\mathcal{B}$, i.e. $(\mathrm{iii})_{n-1}$ and (by induction) (iii) would hold for A, B and $\mathcal{S}'$; repeating the same reasoning one more time, we obtain a covering $\mathcal{C}\subset\mathcal{S}$ such that each $C\in\mathcal{C}$ is disjoint either from A or from B; but it is clear that there is no such covering.

We shall prove in 5.33 a result stronger than 1.6:

1.9 Theorem. *A topological space $(X, \mathcal{T})$ is completely regular if there is a subbase $\mathcal{S}$ for $\mathcal{T}$ satisfying* (iv) *and*

(iii)** *if $A, B\in\mathcal{S}$, $A\cap B=\emptyset$ then there are finite collections $\mathcal{A}$ and $\mathcal{B}$ with $A=\bigcup\mathcal{A}$, $B=\bigcup\mathcal{B}$, and for each pair $A'\in\mathcal{A}$, $B'\in\mathcal{B}$, there is a finite covering $\mathcal{C}\subset\mathcal{S}$ of X such that for each $C\in\mathcal{C}$, there is a $D\in\mathcal{S}$ disjoint from C and containing either A' or B'.*

1.10 Remarks. a) 1.7 a) holds for (iii)**, too.

b) $(\mathrm{iii})^*\Rightarrow(\mathrm{iii})^{**}$ $(D=A''$ or $D=B'')$.

1.11 Example. Theorem 1.9 is strictly stronger than Theorem 1.6. Let

$$X = \{x_i, y_i: i = 0, 1, 2, 3\}\cup\{z_{ij}: i, j = 1, 2, 3\},$$

$$\mathcal{S} = \{\{x_i: 0\leq i\leq 3,\ i\neq j\}, \{y_i: 0\leq i\leq 3,\ i\neq j\}: j = 1, 2, 3\}\cup$$
$$\cup\{\{z_{ij}: x_{m(i)}, y_{m(j)}\}, \{z_{ij}, x_{n(i)}, y_{n(j)}\}: i, j = 1, 2, 3\},$$

where

$$\{i, m(i), n(i)\} = \{1, 2, 3\}, \quad m(i) < n(i)$$

hold for $i=1, 2, 3$. $\mathscr{S}$ is a minimal subbase for the discrete topology. Theorem 1.9 can be applied to $\mathscr{S}$ (check that two distinct points are always contained by disjoint elements of $\mathscr{S}$), but it does not satisfy (iii)*.

Indeed, assume that (iii)* holds and let $A=\{x_0, x_1, x_2\}$, $B=\{y_0, y_1, y_2\}$. Take $A'\in\mathscr{A}$ and $B'\in\mathscr{B}$ with $x_0\in A'$ and $y_0\in B'$. Now the sets A'' and B'' belonging to the pair (A', B') have to be of the form $A''=\{x_i: i\neq i_0\}$, $B''=\{y_j: j\neq j_0\}$. If, $z_{i_0 j_0}\in C\in\mathscr{C}$ then C intersects both A'' and B'', a contradiction.

1.12 Let us consider now the problem whether the condition $\mathscr{A}, \mathscr{B}\subset\mathscr{C}$ could be dropped from (iii) in Theorem 1.1. The answer is negative; not even the following can be substituted for (iii):

(iii)*** if $A, B\in\mathscr{S}$, $A\cap B=\emptyset$ then there are finite collections $\mathscr{A}, \mathscr{B}$ with $A=\bigcup\mathscr{A}$, $B=\bigcup\mathscr{B}$ such that each element of $\mathscr{A}\cup\mathscr{B}$ is the intersection of a finite subfamily of $\mathscr{S}$, and for each pair $A'\in\mathscr{A}$, $B'\in\mathscr{B}$, there is a finite covering $\mathscr{C}\subset\mathscr{S}$ of X such that either $A'\cap C=\emptyset$ or $B'\cap C=\emptyset$ whenever $C\in\mathscr{C}$.

(See Example 1.14.)

1.13 Remarks. a) 1.7 a) holds for (iii)***, too (if $A\subset\bigcup\mathscr{A}$ then replace each $A'\in\mathscr{A}$ by $A'\cap A$).

b) (iii)**$\Rightarrow$(iii)*** (given $A'\in\mathscr{A}$, take $\mathscr{C}(B')$ for each $B'\in\mathscr{B}$; if $C\in\mathscr{C}(B')$ then let $D(B', C)\in\mathscr{S}$ be the set for which we know that $A'\subset D(B', C)$ or $B'\subset D(B', C)$; put now

$$A'^* = A\cap\bigcap\{D(B', C)\colon B'\in\mathscr{B}, C\in\mathscr{C}(B'), A'\subset D(B', C)\};$$

replace each A' by A'^*, and each B' by an analogously defined B'^*).

c) If we drop $\mathscr{A}, \mathscr{B}\subset\mathscr{S}$ from $(\mathrm{iii})_n$ $(n\geqq 2)$ [but keep $(\mathrm{iii})_1=(\mathrm{iii})$ within $(\mathrm{iii})_2$ in its original form] then the resulting condition is equivalent to (iii)*.

1.14 Example. Let $X=\mathbf{R}^2\cup\{\omega\}$; Z, N and I^+ the set of the integers, the positive integers and the positive irrationals, respectively; $N_1=N\setminus\{1\}$,

$$p_{kij} = k+1/i+1/j, \quad q_{kij} = k+1/(i-1)-1/j \qquad (k\in Z,\ i\in N_1, j\in N, j > 2i(i-1)).$$

It is easy to check that the systems $\mathscr{N}(z)$ $(z\in X)$ defined below are neighbourhood bases for a topology $\mathscr{T}$ on X:

$$\mathscr{N}(z) = \begin{cases} \{\{\omega\}\cup]\leftarrow, k]\colon k\in Z\} & \text{if } z = \omega, \\ \{]x-t, x+t[\times\{y\}\colon t\in I^+\} & \text{if } z = (x, y), y \neq \emptyset, x = k+1/i, \\ & \qquad k\in Z,\ i\in N, \\ \{\{z\}\cup((\{p_{kij}, q_{kij}\}\times\mathbf{R})\setminus F)\colon F \text{ is finite}\} & \text{if } z = p_{kij}, \\ \{\{z\}\} & \text{otherwise.} \end{cases}$$

The elements of $\mathscr{N}(z)$ are closed; they are open as well if $z\neq\omega$. A standard argument (see e.g. [45] Example 94) gives that if f is a continuous real function on X then $f=f(\omega)$ on $Z\times\mathbf{R}$, excepting a countable set; therefore $\mathscr{T}$ is not completely regular.

Define now

$$\mathscr{S} = \{X\setminus U\colon\ U\in\mathscr{N}(z),\ z\in\mathbf{R}^2\}\cup$$
$$\cup\{[k+r/i,\ \rightarrow[\times\mathbf{R}\colon\ k\in Z,\ i\in N_1\}\cup$$
$$\cup\{\{\omega\}\cup(]\leftarrow, k+1/i]\times\mathbf{R})\cup(\mathbf{R}\times H)\colon\ k\in Z,\ i\in N_1,\ H=\{1\}\ \text{ or }\ H=\mathbf{R}\setminus\{1\}\}\cup$$
$$\cup\bigcup\{\mathscr{N}(z)\colon\ z\in X\}.$$

$\mathscr{S}$ is a subbase for $\mathscr{T}$ (in fact, the sets in the first and the second line of the definition of $\mathscr{S}$ form a base, and the other elements of $\mathscr{S}$ are closed, too). $\mathscr{S}$ satisfies (iii)*** and (iv). Indeed, the only non-trivial case in (iii)*** is

$$A = \{\omega\}\cup(]\leftarrow, k]\times\mathbf{R}), \qquad B = [k_1+1/i,\ \rightarrow[\times R,$$

where $k, k_1\in Z$, $k\leq k_1$, $i\in N_1$. Take now $\mathscr{A}=\{A\}$, $\mathscr{B}=\{B\cap S_1, B\cap S_2\}$, where

$$S_m = \{\omega\}\cup(]\leftarrow, k+1/(i+1)]\times\mathbf{R})\cup(\mathbf{R}\times H_m) \qquad (m=1,2),$$

$H_1=\{1\}$, $H_2=\mathbf{R}\setminus\{1\}$. For A and $B\cap S_m$, consider the following covering of X:

$$\{S_n, [k+1/(i+1),\ \rightarrow[\times\mathbf{R}\},$$

where $\{m, n\}=\{1, 2\}$.

§ 2. Preproximities

2.0 According to [23], a pair $(\mathscr{S}, r)$ is a *preproximity* on the topological space $(X, \mathscr{T})$ if $\mathscr{S}$ is a subbase for $\mathscr{T}$, r is a symmetrical binary relation on $\mathscr{S}$, and the following conditions are fulfilled:

(1) $\emptyset\ r\ X$;

(2) $A\ r\ B \Rightarrow A\cap B=\emptyset$;

(3) if $A\ r\ B$ then there are finite systems $\mathscr{A}$ and $\mathscr{B}$ such that $A\subset\bigcup\mathscr{A}$, $B\subset\bigcup\mathscr{B}$, and for any pair $A'\in\mathscr{A}$, $B'\in\mathscr{B}$, $A'\ r\ B'$ holds, and there is a finite covering $\mathscr{C}=\mathscr{C}(A', B')$ of X such that either $A'\ r\ C$ or $B'\ r\ C$ whenever $C\in\mathscr{C}$;

(4) if $B\in\mathscr{S}$, $x\in X\setminus B$ then there is a finite system $\mathscr{B}$ such that $B\subset\bigcup\mathscr{B}$, and for each $B'\in\mathscr{B}$, there is an $A'=A'(B')$ with $x\in A'$, $A'\ r\ B'$.

Replace now $\mathscr{S}$ by $\mathscr{S}_1=\operatorname{dom} r$; then conditions (1) to (4) remain valid, since only (4) depends on $\mathscr{S}$, and $\mathscr{S}_1\subset\mathscr{S}$. $\mathscr{S}_1$ is also a subbase for $\mathscr{T}$, since if $B\in\mathscr{S}$, $x\notin B$ then, according to (4) and (2), there is a finite collection $\mathscr{B}\subset\mathscr{S}$ with $x\notin\bigcup\mathscr{B}\supset B$.

Consequently, $\mathscr{S}$ plays no role in the definition; so let us agree that the relation r itself is called a *preproximity* on the set X if it is a symmetrical relation between subsets of X satisfying (1), (2), (3) and

(4*) if $B\in\operatorname{dom} r$, $x\in X\setminus B$ then there is a finite system $\mathscr{B}$ such that $B\subset\bigcup\mathscr{B}$, and for each $B'\in\mathscr{B}$, there is an $A'=A'(B')$ with $x\in A'$, $A'\ r\ B'$.

The *topology induced by r* is the one for which $\operatorname{dom} r$ is a subbase.

2.1 THEOREM [23]. *If r is a preproximity on the set X then there is a coarsest proximity*[1] *δ on X with $r\subset\bar{\delta}$. r and δ induce the same topology.*

[1] Recall that in the case of proximities, the notions *finer* and *coarser* are to be applied to the relation "far" ($\bar{\delta}$).

Conversely, if δ is a proximity then there is a finest preproximity $r \subset \bar{\delta}$. Again, r and δ induce the same topology.

This theorem will be a corollary to Theorem 5.10.

2.2 REMARKS. a) As observed in [26], preproximities can be regarded as subbases for proximities. See also [43] for a related notion.

b) Theorems 1.0 and 1.1 as well as many other characterizations of complete regularity ([44], [10], [47], [50], [6], [1], [32], [11]) can be deduced from Theorem 2.1; see in [23].

c) The word "pre-proximity" is used in a different sense in [36]: it means the same as "d-proximity" (4.6) in our paper.

§ 3. Operations on disjointness relations

3.0 A relation r is a *disjointness relation* (on the set X) if

P0. $A\, r\, B \Rightarrow A \subset X,\ B \subset X$;

P1. $\emptyset\, r\, X,\ X\, r\, \emptyset$;

P2. $A\, r\, B \Rightarrow A \cap B = \emptyset$.

X, the *fundamental set* of r, is uniquely determined by r ($X = \bigcup \operatorname{dom} r = \bigcup \operatorname{ran} r$), and it will be denoted by fund r. Observe that a disjointness relation is not required to be symmetrical.

An *operation on disjointness relations* is a function a defined on the class of disjointness relations, assigning to each r a disjointness relation r^a with fund $r^a =$ fund r. The word *operation* in itself will always mean an operation on disjointness relations. (This convention does not apply to the expression "binary operation".)

The relation $<$ between operations is defined as follows: $a < b$ iff for each disjointness relation r, $r^a \subset r^b$.

3.1 LEMMA. *$<$ is a partial order in the sense "smaller than or equal to" (i.e. it is transitive, and $a = b$ iff $a < b$ and $a > b$).* $\square$

3.2 DEFINITION. For operations a and b, the operations $a \wedge b$ and $a \vee b$ are defined by

$$r^{a \wedge b} = r^a \cap r^b, \qquad r^{a \vee b} = r^a \cup r^b.$$

3.3. LEMMA. *$a \wedge b$ ($a \vee b$) is the infimum (supremum) of the operations a and b with respect to the partial order $<$.* $\square$

3.4 DEFINITION. For operations a and b, the *product* ab of a and b is defined by $r^{ab} = (r^a)^b$.

3.5 LEMMA. *Each of the binary operations $\wedge$, $\vee$ and the multiplication is associative; $\wedge$ and $\vee$ are commutative.* $\square$

3.6 Definition. An operation a is *monotone* if $r_1 \subset r_2$ and fund $r_1 = \text{fund}\, r_2$ imply $r_1^a \subset r_2^a$.

3.7 Lemma. *If a and b are monotone operations then so are $a \wedge b$, $a \vee b$ and ab.* □

3.8 Lemma. *Let $a_1 < b_1$, $a_2 < b_2$, and assume that either a_2 or b_2 is monotone. Then $a_1 a_2 < b_1 b_2$.*

Proof. 1° Assume that a_2 is monotone. $a_1 < b_1$ means $r^{a_1} \subset r^{b_1}$, thus $r^{a_1 a_2} \subset r^{b_1 a_2}$; from $a_2 < b_2$ we have $r^{b_1 a_2} \subset r^{b_1 b_2}$, therefore $r^{a_1 a_2} \subset r^{b_1 b_2}$, i.e. $a_1 a_2 < b_1 b_2$.
2° If b_2 is monotone then $r^{a_1 a_2} \subset r^{a_1 b_2} \subset r^{b_1 b_2}$ can be proved in the same way. □

3.9 Definition. A *formula* is a finite formula built up from variables and the binary operations $\wedge$, $\vee$ and the multiplication.

3.10 Lemma. *Let* $\mathbf{F}$ *be a formula of n variables and assume that the operations $a_1, \ldots, a_n$ are monotone. Then* $\mathbf{F}(a_1, \ldots, a_n)$ *is monotone, too.*

Proof. 3.7 and induction on the length of $\mathbf{F}$. □

3.11 Lemma. *Let* $\mathbf{F}$ *be a formula of n variables and assume that a_i and b_i are monotone operations, $a_i < b_i$ $(i = 1, \ldots, n)$. Then*

$$\mathbf{F}(a_1, \ldots, a_n) < \mathbf{F}(b_1, \ldots, b_n).$$

Proof. If $a_1 < b_1$ and $a_2 < b_2$ then clearly $a_1 \wedge a_2 < b_1 \wedge b_2$ and $a_1 \vee a_2 < b_1 \vee b_2$. Thus 3.8 and 3.10 make a proof by induction possible. □

3.12 Definition. h is the identity operation, i.e. $r^h = r$ for each disjointness relation r. The operation -1 is defined by $A\, r^{-1}\, B$ iff $B\, r\, A$.

3.13 Definition. The operations d_1 and d_2 are defined as follows:
a) $A\, r^{d_1}\, B$ iff there is a set $A' \supset A$ with $A'\, r\, B$;
b) $A\, r^{d_2}\, B$ iff there is a set $B' \supset B$ with $A\, r\, B'$.

3.14 Lemma. a) *d_1 and d_2 are monotone operations;*
b) $d_1 d_1 = d_1$, $d_2 d_2 = d_2$;
c) $h < d_1$, $h < d_2$;
d) $d_1 d_2 = d_2 d_1$. □

3.15 Definition. $d = d_1 d_2$.

3.16 Lemma. a) *d is a monotone operation;*
b) $dd = d$;
c) $d_1 < d$, $d_2 < d$. □

3.17 Definition. The operations q_1, q_2 and q are defined as follows:
a) $A\, r^{q_1}\, B$ iff there is a finite collection $\mathscr{A} \neq \emptyset$ such that $A = \cup \mathscr{A}$, and $A'\, r\, B$ holds for each $A' \in \mathscr{A}$;
b) $A\, r^{q_2}\, B$ iff there is a finite collection $\mathscr{B} \neq \emptyset$ such that $B = \cup \mathscr{B}$, and $A\, r\, B'$ holds for each $B' \in \mathscr{B}$;
c) $A\, r^q\, B$ iff there are finite non-empty collections $\mathscr{A}$ and $\mathscr{B}$ such that $A = \cup \mathscr{A}$, $B = \cup \mathscr{B}$, and $A'\, r\, B'$ whenever $A' \in \mathscr{A}$ and $B' \in \mathscr{B}$.

3.18 LEMMA. a) *q_1, q_2 and q are monotone operations;*
b) $h < q_1 < q$, $h < q_2 < q$;
c) $q < q_1 q_2 < d_1 q$, $q < q_2 q_1 < d_2 q$.

PROOF. We shall only show that $q_1 q_2 < d_1 q$; the proof of $q_2 q_1 < d_2 q$ is analogous, the other statements are evident.

If $A\, r^{q_1 q_2} B$ then there are B_i $(1 \leqq i \leqq n)$ such that $B = \bigcup_j B_i$, and $A\, r^{q_1} B_i$ for each i; thus there are sets A_{ij} $(1 \leqq i \leqq n,\ 1 \leqq j \leqq m_i)$ such that $A = \bigcup_i A_{ij}$ $(1 \leqq i \leqq n)$, and $A_{ij}\, r\, B_i$ for any pair (i, j). Put

$$A(x) = \bigcap \{A_{ij} \colon x \in A_{ij}\} \qquad (x \in A).$$

Given $x \in A$ and $1 \leqq i \leqq n$, there is a j with $x \in A_{ij}$, so $A(x)\, r^{d_1} B_i$ follows from $A(x) \subset A_{ij}$ and $A_{ij}\, r\, B_i$. Now the finite collections $\{A(x) \colon x \in A\}$ (or $\{\emptyset\}$ if $A = \emptyset$) and $\{B_i \colon 1 \leqq i \leqq n\}$ show that $A\, r^{d_1 q} B$. □

3.19 REMARKS. a) Notions closely related to the operation q are often denoted by the same symbol, see e.g. [5] Chapter 3 and [9] § 3.
b) $q \neq q_1 q_2$.

3.20 LEMMA. a) $q_1 q_1 = q_1$, $q_2 q_2 = q_2$;
b) $qq < dq$.

PROOF. b) By 3.8, 3.7, 3.18 a), 3.18 c), 3.20 a), 3.14 a) and 3.15, $qq < q_1 q_2 q_2 q_1 = q_1 q_2 q_1 < d_1 q q_1 < d_1 q_2 q_1 q_1 = d_1 q_2 q_1 < d_1 d_2 q = dq$. □

3.21 LEMMA. a) $q_1 d_1 = d_1 q_1$, $q_2 d_2 = d_2 q_2$;
b) $q_1 d_2 < d_2 q_1$, $q_2 d_1 < d_1 q_2$;
c) $qd < dq$.

PROOF. a) $A\, r^{q_1 d_1} B$ and $A\, r^{d_1 q_1} B$ both mean that there is a finite collection $\mathscr{A}$ such that $A \subset \bigcup \mathscr{A}$, and $A'\, r\, B$ holds for each $A' \in \mathscr{A}$.
b) Evident.
c) By 3.18 c), 3.15, 3.14 c), 3.21 a), 3.21 b) and 3.14 b), $qd < q_1 q_2 d_2 d_1 < d_2 d_1 q_1 q_2 < d_2 d_1 d_1 q = d_2 d_1 q = dq$. □

3.22 LEMMA. *If $a = \mathbf{F}(h, d_1, d_2, d, q_1, q_2, q)$ with an arbitrary formula $\mathbf{F}$ then $h < a < dq$.*

PROOF by induction. 1° If a is equal to any of the operations $h, d_1, d_2, d, q_1, q_2, q$ then $h < a$ by 3.14 c), 3.16 c) and 3.18 b). On the other hand, the same lemmas and 3.8 give that $a < dq$ (e.g. $d_1 < dq$ follows from $d_1 = d_1 h$, $d_1 < d$ and $h < q$).

2° Assume $h < a_i < dq$ $(i = 1, 2)$. Now $h = hh < a_1 a_2$ (3.8) and $a_1 a_2 < dqdq < ddqq < dddq = dq$ (3.21 c), 3.20 b) and 3.16 b)). The inequalities $h < a_1 \wedge a_2 < dq$ and $h < a_1 \vee a_2 < dq$ are evident (3.3). □

3.23 DEFINITION. The operations c_0 and c are defined as follows:
a) $A\, r^{c_0} B$ iff there are A' and B' such that $A' \cup B' = \operatorname{fund} r$, $A'\, r\, B$, $A\, r\, B'$;
b) $A\, r^{c} B$ iff $A, B \subset \operatorname{fund} r$, and there is a finite collection $\mathscr{C}$ such that $\bigcup \mathscr{C} = \operatorname{fund} r$, and for each $C \in \mathscr{C}$, either $A\, r\, C$ or $C\, r\, B$.

3.24 NOTATION. If $\mathscr{C}$ is chosen according to 3.23 b) then let

$$\mathscr{C}^1 = \mathscr{C}^1(A, B, r) = \{C\in\mathscr{C}:\ C\ r\ B\},$$
$$\mathscr{C}^2 = \mathscr{C}^2(A, B, r) = \{C\in\mathscr{C}:\ A\ r\ C\}.$$

3.25 LEMMA. a) *If* $A\ r^{c_0}\ B$ *then* $A\subset A'$ *and* $B\subset B'$ *hold for the sets* A' *and* B' *in* 3.23 a).

b) *If* $A\ r^c\ B$ *then* $A\subset \bigcup\mathscr{C}^1$, $B\subset\bigcup\mathscr{C}^2$, $\bigcup\mathscr{C}^1\cup\bigcup\mathscr{C}^2=\text{fund}\ r$ *hold for the collection* $\mathscr{C}$ *in* 3.23 b).

PROOF. Use Axiom P2. □

3.26 LEMMA. a) c_0 *and* c *are monotone operations;*

b) $c_0<c<qc_0d$;

c) *if* $A\subset\text{fund}\ r$ *then* $A\ r^c\ \emptyset$, $\emptyset\ r^c\ A$;

d) $c_0c_0<c_0$, $cc<c$.

PROOF. a) Let r be a disjointness relation, $X=\text{fund}\ r$. In order to prove that c is an operation, we have to show that r^c satisfies P0, P1 and P2 with the same X. P0 is contained in the definition; P1 can be proved by choosing $\mathscr{C}=\{X\}$; P2 follows from $A\subset\bigcup\mathscr{C}^1$ and $C\cap B=\emptyset$ $(C\in\mathscr{C}^1)$. The case of c_0 is similar (to show P1, take $\{A', B'\}=\{\emptyset, X\}$). c and c_0 are evidently monotone.

b) 1° To prove $c_0<c$, take $\mathscr{C}=\{A', B'\}$.

2° If $A\ r^c\ B$ and $A\neq\emptyset\neq B$ then, putting $A'=\bigcup\mathscr{C}^1$ and $B'=\bigcup\mathscr{C}^2$, we have $A'\ r^q\ B$, $A\ r^q\ B'$ and $A'\cup B'=\text{fund}\ r$, thus $A\ r^{qc_0}\ B$ and $A\ r^{qc_0d}\ B$. If $A=\emptyset$ or $B=\emptyset$ then $A\ r^{qc_0d}\ B$ is evident.

c) Take $\mathscr{C}=\{X\}$.

d) 1° Assume $A\ r^{cc}\ B$. Choose a finite covering $\mathscr{C}$ of fund $r^c=\text{fund}\ r$ such that for each $C\in\mathscr{C}$, $A\ r^c\ C$ or $C\ r^c\ B$. If $C\in\mathscr{C}^1=\mathscr{C}^1(A, B, r^c)$ then choose a finite covering $\mathscr{D}_C$ such that either $C\ r\ D$ or $D\ r\ B$ whenever $D\in\mathscr{D}_C$; similarly, if $C\in\mathscr{C}^2$ then let $\mathscr{E}_C$ be a finite covering such that either $A\ r\ E$ or $E\ r\ C$ whenever $E\in\mathscr{E}_C$. Now the covering

$$\mathscr{F} = \bigcup\{\mathscr{D}_C^1(C, B, r):\ C\in\mathscr{C}^1\}\cup\bigcup\{\mathscr{E}_C^2(A, C, r):\ C\in\mathscr{C}^2\}$$

shows that $A\ r^c\ B$ (to prove that $\mathscr{F}$ is a covering, apply 3.25 b) first to $\mathscr{C}$ then to each $\mathscr{D}_C$ and $\mathscr{E}_C$).

2° The proof of $c_0c_0<c_0$ is similar. □

3.27 REMARK. Let us consider the following modifications of q and c:

a) $A\ r^{q^*}\ B$ iff $A, B\subset\text{fund}\ r$ and there are (possibly empty) finite collections $\mathscr{A}$ and $\mathscr{B}$ such that $A=\bigcup\mathscr{A}$, $B=\bigcup\mathscr{B}$, and $A'\ r\ B'$ whenever $A'\in\mathscr{A}$ and $B'\in\mathscr{B}$;

b) $A\ r^{c^*}\ B$ iff there are finite non-empty collections $\mathscr{A}$ and $\mathscr{B}$ such that $\bigcup\mathscr{A}\cup\bigcup\mathscr{B}=\text{fund}\ r$, and $A'\ r\ B$ if $A'\in\mathscr{A}$, $A\ r\ B'$ if $B'\in\mathscr{B}$.

For non-empty sets A and B, $A\ r^{q^*}\ B$ iff $A\ r^q\ B$, $A\ r^{c^*}\ B$ iff $A\ r^c\ B$. Clearly, $c_0<c^*<c$, $q<q^*<qd$. In the more important statements, q and q^*, respectively c and c^* are interchangeable. 3.26 b) holds in a stronger form for c^* or q^*: $c<q^*c_0$, $c^*<qc_0$, $c^*<q^*c_0$. We have preferred q to q^* on account of 5.31, and c to c^* because its definition looks somewhat simpler.

3.28 LEMMA. $c<q_1d$, $c<q_2d$.

PROOF. Assume $A\ r^c\ B$ and choose $\mathscr{C}$ according to 3.23 b). If $\mathscr{C}^1=\emptyset$ then $A=\emptyset$, thus $A\ r^{q_1d}\ B$; otherwise $(\bigcup\mathscr{C}^1)\ r^{q_1}\ B$, thus $A\ r^{q_1d}\ B$ again. □

3.29 LEMMA. $cd<dc$, $c_0d=dc_0$.

Proof. 1° Assume $A_1\ r^{cd}\ B_1$ and take $A\supset A_1$, $B\supset B_1$ with $A\ r^c\ B$. Choose $\mathscr{C}$ according to 3.23 b). Then $A\ r\ C$ or $C\ r\ B$ holds for each $C\in\mathscr{C}$, thus either $A_1\ r^d\ C$ or $C\ r^d\ B_1$, i.e. the covering $\mathscr{C}$ shows that $A_1\ r^{dc}\ B_1$.

2° The proof of $c_0d<dc_0$ is similar.

3° Assume $A\ r^{dc_0}\ B$ and take A', B' according to 3.23 a). $A\ r^d\ B'$ implies the existence of $A_1\supset A$ and $B''\supset B'$ with $A_1\ r\ B''$. Similarly, there are $A''\supset A'$ and $B_1\supset B$ with $A''\ r\ B_1$. Now the pair (A'', B'') shows that $A_1\ r^{c_0}\ B_1$, therefore $A\ r^{c_0d}\ B$. □

3.30 LEMMA. $qc_0<cq$.

PROOF. Assume $A\ r^{qc_0}\ B$. Then there are E and F with $E\cup F=\text{fund}\ r$, $A\ r^q\ F$, $E\ r^q\ B$, i.e. there are finite non-empty collections $\mathscr{A}, \mathscr{B}, \mathscr{E}, \mathscr{F}$ with $A=\bigcup\mathscr{A}$, $B=\bigcup\mathscr{B}$ $E=\bigcup\mathscr{E}$, $F=\bigcup\mathscr{F}$ such that $A'\ r\ F'$ if $A'\in\mathscr{A}$, $F'\in\mathscr{F}$, and $E'r\ B'$ if $E'\in\mathscr{E}$, $B'\in\mathscr{B}$. Now the finite covering $\mathscr{E}\cup\mathscr{F}$ shows that $A'\ r^c\ B'$ whenever $A'\in\mathscr{A}$ and $B'\in\mathscr{B}$; therefore $A\ r^{cq}\ B$. □

3.31 LEMMA. $cq<dqc$.

PROOF. Assume $A\ r^{cq}\ B$ and let $A=\bigcup\limits_i A_i$, $B=\bigcup\limits_j B_j$, $A_i\ r^c\ B_j$. Choose coverings $\mathscr{C}_{ij}=\{C_{ijk}\}$ such that $A_i\ r\ C_{ijk}$ or $C_{ijk}\ r\ B_j$ for any triplet (i,j,k). Put

$$C(x) = \bigcap\{C_{ijk}\colon x\in C_{ijk}\}, \qquad \mathscr{C} = \{C(x)\colon x\in\text{fund}\ r\}.$$

$\mathscr{C}$ is evidently a finite covering.

In order to prove $A\ r^{dqc}\ B$, it is enough to show that for any $x\in\text{fund}\ r$, either $A\ r^{dq}\ C(x)$ or $C(x)\ r^{dq}\ B$; so it is enough to show that either $A_i\ r^d\ C(x)$ for each i, or $C(x)\ r^d\ B_j$ for each j. Assume the contrary, and choose i, j and k such that $A_i\ \overline{r^d}\ C(x)$, $C(x)\ \overline{r^d}\ B_j$ and $x\in C_{ijk}$; now $A_i\ \bar{r}\ C_{ijk}$ and $C_{ijk}\ \bar{r}\ B_j$, a contradiction. □

3.32 LEMMA. $dcq=dqc=dqc_0$.

PROOF.

$$dcq \underset{1}{<} ddqc \underset{2}{=} dqc \underset{3}{<} dqqc_0d \underset{4}{=} dqqdc_0 \underset{5}{<} dqc_0 \underset{6}{<} dcq$$

[1: 3.31; 2: 3.16 b); 3: 3.26 b); 4: 3.29; 5: 3.22; 6: 3.30]. □

3.33 DEFINITION. $e=dcq$.

3.34 EXAMPLE. We are going to show that none of the operations dc_0q, $\bar{c}dq$, $\bar{c}qd$, $q\bar{c}d$ and $qd\bar{c}$ ($\bar{c}=c$ or c_0) is equal to e. Since

$$c_0qd \underset{1}{<} cqd \underset{2}{<} cdq \underset{3}{<} dcq = e,$$

$$dc_0q \underset{4}{=} c_0dq \underset{1}{<} cdq < e,$$

$$qdc_0 \underset{4}{=} qc_0d \underset{1}{<} qcd \underset{3}{<} qdc \underset{2}{<} dqc \underset{5}{=} e$$

[1 : 3.26 b); 2: 3.21 c); 3 and 4: 3.29; 5: 3.32], it is enough to produce a disjointness relation r such that $r^{qdc} \neq r^e \neq r^{cdq}$. The example will be symmetrical, too.

Let r be the coarsest symmetrical disjointness relation on $\{1, 2, 3, 4, 5\}$ for which

$$\{i, j\}\, r\, \{k\} \quad (1 \leq i \leq 5,\ j = i+1 \ (\mathrm{mod}\ 5),\ k = j+1 \ (\mathrm{mod}\ 5)).$$

Check that if $A=\{1\}$, $B=\{2, 3, 4\}$ then $A\, r^e\, B$, but $A\, \overline{r^{qdc}}\, B$ and $A\, \overline{r^{cdq}}\, B$.

3.35 LEMMA. *If* $a=\mathbf{F}(h, d_1, d_2, d, q_1, q_2, q, c_0, c)$ *(***F** *an arbitrary formula) then* $a<dq$.

PROOF. By 3.26 b), 3.28, 3.11 and 3.22,

$$a < \mathbf{F}(h, d_1, d_2, d, q_1, q_2, q, q_1 d, q_1 d) < dq. \quad \square$$

3.36 LEMMA. *If* $\tilde{c} \in \{a_1, \ldots, a_n\} \subset \{h, d_1, d_2, d, q_1, q_2, q, c_0, c\}$ $(\tilde{c}=c$ or $c_0)$ *then* $a_1 \ldots a_n < e$.

PROOF. By 3.35, $a_1 \ldots a_n < dq\tilde{c}dq$, so

$$a_1 \ldots a_n \underset{1}{<} dqcdq \underset{2}{<} dqdcq \underset{3}{=} dqdqc \underset{4}{<} dqc \underset{5}{=} e$$

[1 : 3.26 b); 2: 3.29; 3: 3.32; 4: 3.35; 5: 3.32]. $\square$

3.37 LEMMA. *If* $a=\mathbf{F}(h, d_1, d_2, d, q_1, q_2, q, c_0, c)$ *(***F** *an arbitrary formula) then* $h<a$ *or* $a<e$.

PROOF by induction. 1° If $a=h, d_1, d_2, d, q_1, q_2$ or q then $h<a$ by 3.22; if $a=c_0$ or c then $a<e$ by 3.36.

2° Assume $h<a_1$ or $a_1<e$, and $h<a_2$ or $a_2<e$. If $h<a_1$ and $h<a_2$ then $h<a_1 a_2$. If $a_1<e$ or $a_2<e$ then $a_1 a_2<edq$ or dqe (3.35), thus $a_1 a_2<e$ (3.36). If $h<a_1$ and $h<a_2$ then $h<a_1 \wedge a_2$; otherwise $a_1 \wedge a_2<e$. If $a_1<e$ and $a_2<e$ then $a_1 \vee a_2<e$; otherwise $h<a_1 \vee a_2$. $\square$

3.38 LEMMA. $e=dcdq$. $\square$

3.39 LEMMA (essentially [23] Lemma 2.2). *$A\, r^{dq}\, B$ iff $A, B \subset \mathrm{fund}\, r$ and there is a finite collection $\mathscr{F}$ such that for any $x \in A$ and $y \in B$, there are $A', B' \in \mathscr{F}$ with $x \in A'$, $y \in B'$, $A'\, r\, B'$.*

PROOF. 1° If $A\, r^{dq}\, B$ then there are sets A_i, B_j, A_i', B_j' $(1 \leq i \leq n,\ 1 \leq j \leq m)$ such that $A=\bigcup_i A_i$, $B=\bigcup_j B_j$, $A_i' \supset A_i$, $B_j' \supset B_j$, $A_i'\, r\, B_j'$. Now

$$\mathscr{F} = \{A_i' \colon 1 \leq i \leq n\} \cup \{B_j' \colon 1 \leq j \leq m\}$$

will do.

2° Assume there is an $\mathscr{F}$ as described in the lemma. If $A=\emptyset$ or $B=\emptyset$ then $A\, r^d\, B$ by P1, thus $A\, r^{dq}\, B$. So let $A \neq \emptyset \neq B$. If $z \in A \cup B$ then there is at least one $F \in \mathscr{F}$ containing z (e.g. if $z \in A$ then pick $y \in B$, choose $A', B' \in \mathscr{F}$ with $z \in A'$, $y \in B'$ and put $F=A'$), thus we can define

$$F(z) = \bigcap\{F \in \mathscr{F} \colon z \in F\} \qquad (z \in A \cup B).$$

Put

$$\mathscr{A} = \{A \cap F(x) \colon x \in A\}, \qquad \mathscr{B} = \{B \cap F(y) \colon y \in B\}.$$

For $A_0=F(x)\cap A$ and $B_0=F(y)\cap B$, there are $A', B'\in\mathscr{F}$ with $x\in A'$, $y\in B'$, $A'\ r\ B'$. By the definition of $F(x)$ and $F(y)$, we have $A_0\subset A'$ and $B_0\subset B'$, therefore $A_0\ r^d B_0$ and $A\ r^{dq}\ B$. □

3.40 LEMMA. *$A\ r^e\ B$ iff $A, B\subset \operatorname{fund} r$, and there is a finite collection $\mathscr{E}$ such that for any triplet $x\in A$, $y\in B$, $z\in\operatorname{fund} r$, there are $C=C(x, y, z)\in\mathscr{E}$ and $D=D(x, y, z)\in\mathscr{E}$ with $z\in C$ and either $x\in D$, $D\ r\ C$ or $y\in D$, $C\ r\ D$.*

PROOF. 1° Assume that $A\neq\emptyset\neq B$ and there is such an $\mathscr{E}$. Let $\mathscr{F}$ denote the family of all intersections of subsystems of $\mathscr{E}$. Fixing $x\in A$ and $y\in B$, let A' and B' be the smallest elements of $\mathscr{F}$ containing x, respectively y; furthermore, put

$$\mathscr{C} = \{C(x, y, z)\colon\ z\in\operatorname{fund} r\}.$$

Now $\mathscr{C}$ is a covering such that for each $C\in\mathscr{C}$, either $A'\ r^d\ C$ or $C\ r^d\ B'$; therefore $A'\ r^{dc}\ B'$, $A\ r^{dcdq}\ B$ (3.39) and $A\ r^e\ B$ (3.38).

2° Conversely, if $A\ r^e\ B$ then $A\ r^{dqc_0}\ B$ (3.32), so there are A' and B' such that $A\ r^{dq}\ B'$, $A'\ r^{dq}\ B$ and $A'\cup B'=\operatorname{fund} r$. Choose $\mathscr{F}_1$ for (A, B') and $\mathscr{F}_2$ for (A', B) according to 3.39. Now $\mathscr{E}=\mathscr{F}_1\cup\mathscr{F}_2$ will do. □

3.41 DEFINITION. For an operation a, $a^{-1}=(-1)a(-1)$. a is *symmetrical* if $a^{-1}=a$.

3.42 LEMMA. a) *$(a^{-1})^{-1}=a$. a is symmetrical iff $(-1)a=a(-1)$.*

b) *If the disjointness relation r and the operation a are both symmetrical then so is r^a.* □

3.34 LEMMA. *If $a_1, \dots, a_n$ are symmetrical operations then so is $\mathbf{F}(a_1, \dots, a_n)$, where $\mathbf{F}$ is an arbitrary formula. h, d, q, c_0 and c are symmetrical. $d_1^{-1}=d_2$, $q_1^{-1}=q_2$.* □

3.44 NOTATION. If $\mathscr{S}$ is a family of subsets of a set X then the relation $r_{\mathscr{S}}=r_{\mathscr{S},X}$ is defined by

$$A\ r_{\mathscr{S}} B \Leftrightarrow A\cap B=\emptyset,\ A, B\in\mathscr{S}\cup\{\emptyset, X\}.$$

3.45 LEMMA. *$r_{\mathscr{S},X}$ is a symmetrical disjointness relation on X.* $\operatorname{dom} r_{\mathscr{S}}=\operatorname{ran} r_{\mathscr{S}}=\mathscr{S}\cup\{\emptyset, X\}$. □

3.46 DEFINITION. For a disjointness relation r on X, topologies $\mathscr{T}_r^1$ and $\mathscr{T}_r^2$ on X are defined as follows:

$$A\in\mathscr{T}_r^1 \Leftrightarrow \forall y\in X\setminus A,\ A\ r^{dq}\ \{y\};$$

$$B\in\mathscr{T}_r^2 \Leftrightarrow \forall x\in X\setminus B,\ \{x\}\ r^{dq}\ B.$$

We shall say that the *bitopology* $(\mathscr{T}_r^1, \mathscr{T}_r^2)$ *is induced by* r, or that r is *compatible with the bitopology* $(\mathscr{T}_r^1, \mathscr{T}_r^2)$. If $\mathscr{T}_r^1=\mathscr{T}_r^2=\mathscr{T}_r$ then the *topology* $\mathscr{T}_r$ *is induced by* r, or r is *compatible with the topology* $\mathscr{T}_r$.

3.47 LEMMA. a) *$\mathscr{T}_r^1$ and $\mathscr{T}_r^2$ are indeed topologies.*

b) *If r_1 is finer than r_2 then $\mathscr{T}_{r_1}^i$ is finer than $\mathscr{T}_{r_2}^i$ $(i=1, 2)$.*

c) *$\mathscr{T}_{r^{-1}}^1=\mathscr{T}_r^2$. If r is symmetrical then $\mathscr{T}_r^1=\mathscr{T}_r^2$.*

PROOF. a) $\emptyset, X\in\mathcal{T}_r^1$ is evident. If $A_1, A_2\in\mathcal{T}_r^1$ and $y\notin A_1\cup A_2$ then $A_i\, r^{dq}\{y\}$ $(i=1,2)$, thus $(A_1\cup A_2)\, r^{dqq}\{y\}$, therefore $(A_1\cup A_2)\, r^{dq}\{y\}$ (3.35). If $A_\alpha\in\mathcal{T}_r^1$ $(\alpha\in I)$ and $y\notin A=\bigcap_\alpha A_\alpha$ then take an $\alpha_0\in I$ with $y\notin A_{\alpha_0}$; now $A_{\alpha_0} r^{dq}\{y\}$, thus $A\, r^{dqd}\{y\}$, i.e. $A\, r^{dq}\{y\}$, again by 3.35. □

3.48 LEMMA. $\mathcal{T}_r^i=\mathcal{T}_{r^d}^i=\mathcal{T}_{r^q}^i$ $(i=1,2)$.

PROOF. $dq=ddq=qdq$. □

3.49 LEMMA. a) dom r *is a subbase for* $\mathcal{T}_r^1$ *iff* dom $r\subset\mathcal{T}_r^1$.
b) ran r *is a subbase for* $\mathcal{T}_r^2$ *iff* ran $r\subset\mathcal{T}_r^2$.

PROOF. a) Assume dom $r\subset\mathcal{T}_r^1$, and take an $A\in\mathcal{T}_r^1$, $\emptyset\neq A\neq\operatorname{fund} r$. For each $y\notin A$, we have $A\, r^{dq}\{y\}$, i.e. there are sets $A_i(y)\in\operatorname{dom} r$ $(i=1,\dots,n(y))$ such that $A\subset\bigcup_i A_i(y)=S(y)$ and, by P2, $y\notin S(y)$. Now $A=\bigcap\{S(y)\colon y\notin A\}$, so dom r is a subbase for $\mathcal{T}_r^1$. □

3.50 NOTATION. $\mathrm{cl}_r^i=\mathrm{cl}_{\mathcal{T}_r^i}$ $(i=1,2)$.

3.51 DEFINITION. The operation w is defined by

$$A\, r^w B \Leftrightarrow A\in\mathcal{T}_r^1,\quad B\in\mathcal{T}_r^2,\ A\, rB.$$

3.52 LEMMA. a) *w is a symmetrical monotone operation.*
b) $w<h$.

PROOF. a) w is monotone by 3.47 b). □

3.53 DEFINITION. An *elementary operation* is an operation of the form

$$\mathbf{F}(h, d_1, d_2, d, q_1, q_2, q, c_0, c, w)$$

where $\mathbf{F}$ is a formula.

3.54 LEMMA. *If a is an elementary operation then*
a) $a<dq$;
b) *either* $a<e$ *or there is a positive integer n such that* $w^n<a$.

PROOF. a) By 3.52 b), 3.11 and 3.35:

$$a<\mathbf{F}(h, d_1, d_2, d, q_1, q_2, q, c_0, c, h)<dq.$$

b) (Induction.) 1° If $a=h, d_1, d_2, d, q_1, q_2, q, c_0$ or c then $a<e$ or $w<h<a$ (3.37 and 3.52b)); if $a=w$ then $w<a$.

2° Assume that the statement of the lemma holds for a_1 and a_2. If $a_1a_2\not< e$ then $a_1\not< e$ and $a_2\not< e$ (otherwise 3.54 a) would imply $a_1a_2<edq$ or $a_1a_2<dqe$, i.e. $a_1a_2<e$ by 3.36). Thus $w^{n_i}<a_i$ $(i=1,2)$ and $w^{n_1+n_2}<a_1a_2$.

If $a_1\vee a_2\not< e$ then $a_1\not< e$ or $a_2\not< e$, thus $w^n<a_i$ $(i=1$ or $2)$ and $w^n<a_1\vee a_2$.

If $a_1\wedge a_2\not< e$ then $a_i\not< e$ $(i=1,2)$, thus $w^{n_i}<a_i$ $(i=1,2)$ and, by 3.52 b), $w^n<a_1\wedge a_2$, where $n=\max\{n_1, n_2\}$. □

§ 4. Quasi-proximities

4.0 Proximities and their generalizations can be described in the following terms:
$A\ \delta\ B = B$ is near to A (the orthodox way);
$A\ r\ B = B$ is far from A ([48]);
$A < B = B$ is a proximal neighbourhood of A ([44], [5]).
Here $r = \bar{\delta}$; $A < B$ iff $A\ r\ (\text{fund}\ r \setminus B)$. We shall take r ("far") as the primitive term; then the axioms for a quasi-proximity (cf. [40], [46], [39]) can be given in the following form:

4.1 Definition. A *quasi-proximity* is a disjointness relation r satisfying

Q1. $r^{d_1} \subset r$, $r^{d_2} \subset r$;

Q2. $r^{q_1} \subset r$, $r^{q_2} \subset r$;

Q3. $r \subset r^{c_0}$.

4.2 Remarks. a) A *proximity* is a symmetrical quasi-proximity (cf. [16], [39], [17], [8]).

b) A notion equivalent to a quasi-proximity was first introduced in [3] ("topogenous structure").

4.3 Lemma. a) *For a disjointness relation r, any of the following conditions is equivalent to* Q1: (i) $r^{d_1} = r$, $r^{d_2} = r$; (ii) $r^d \subset r$; (iii) $r^d = r$.

b) *An analogous statement for* Q2, q_1, q_2 *and* q.

Proof. a) 3.14 c), 3.16 c), 3.14 a). b) 3.18. □

4.4 Lemma. *For a disjointness relation r,* Q1+Q2 *is equivalent to any of the following conditions:* (i) $r^{dq} \subset r$; (ii) $r^{dq} = r$; (iii) $r^{qd} \subset r$; (iv) $r^{qd} = r$. □

4.5 Lemma. *For a disjointness relation r satisfying* Q1 *and* Q2, *any of the following conditions is equivalent to* Q3: (i) $r = r^{c_0}$; (ii) $r \subset r^c$; (iii) $r = r^c$; (iv) $r \subset r^e$; (v) $r = r^e$.

Proof. $r^c = r^{dqc} = r^e = r^{dqc_0} = r^{c_0}$, so Q3⇔(ii)⇔(iv) and (i)⇔(iii)⇔(v), (i)⇒Q3 is evident, while Q3⇒(i) follows from $r^{c_0} \subset r^{dq} = r$ (3.35). □

4.6 Definition. Let a be an operation. A disjointness relation r is an *a-proximity* if $r = r^a$.

4.7 Lemma. *If* $\bar{c} = c$ *or* c_0, $\{a_1, a_2, a_3\} = \{\bar{c}, d, q\}$ *then a disjointness relation r is a quasi-proximity iff* $r = r^{a_1 a_2 a_3}$. *In particular, the quasi-proximities are just the e-proximities.*

Proof. 1° If r is a quasi-proximity then $r^{a_i} = r$ $(i = 1, 2, 3)$ by 4.3 and 4.5, therefore $r^{a_1 a_2 a_3} = r$.

2° Conversely, assume $r = r^{a_1 a_2 a_3}$. Then

$$r^d = r^{a_1 a_2 a_3 d} \subset r^{a_1 a_2 a_3} = r,$$

by 3.21 c), 3.29 and 3.16 b); thus Q1 holds, and $r^d = r$ (4.3).

Define i and j by $\tilde{c}=a_i$ and $q=a_j$. If $i<j$ then

$$r^q = r^{a_1a_2a_3q} \underset{1}{\subset} r^{d\tilde{c}qq} \underset{2}{=} r^{\tilde{c}qq} \underset{3}{\subset} r^{\tilde{c}dq} \underset{4}{\subset} r^{d\tilde{c}q} \underset{2}{=} r^{\tilde{c}q} \underset{5}{\subset} r^{a_1a_2a_3} = r$$

[1: 3.21 c) and 3.29; 2: $r^d=r$; 3: 3.20 b); 4: 3.29; 5: $h<d$]. If $j<i$ then

$$r^q = r^{a_1a_2a_3q} \underset{1}{\subset} r^e \underset{2}{=} r^{dq\tilde{c}} \underset{3}{=} r^{q\tilde{c}} \underset{4}{\subset} r^{a_1a_2a_3} = r$$

[1: 3.36; 2: 3.32, 3: $r^d=r$; 4: $h<d$]. Thus Q2 holds (4.3).

Finally, $r=r^{a_1a_2a_3}\subset r^e$ by 3.36, thus Q3 follows from 4.5. □

4.8 LEMMA. *Each e-proximity is a dq-proximity; each dq-proximity is a d-proximity.* □

4.9 REMARK. dq-proximities (respectively, d-proximities) are essentially the same as the topogenous orders (respectively, semi-topogenous orders) introduced in [3], see also [4], [5]. The symmetrical dq-proximities (symmetrical d-proximities) are the proximities in the sense of [2] (the semi-proximities in the sense of [9]).

4.10 LEMMA. a) *If r is a disjointness relation the r^d is the coarsest d-proximity finer than r.*

b) [5] (3.6) *If r is a d-proximity then r^q is the coarsest dq-proximity finer than r.*

c) *If r is a disjointness relation then r^{dq} is the coarsest dq-proximity finer than r.*

PROOF. a) $r\subset r^d$; $(r^d)^d=r^{dd}=r^d$; *if* $r\subset r_1$ and r_1 is a d-proximity then $r^d\subset r_1^d=r_1$.

b) $r\subset r^q$; $(r^q)^{dq}=r^{qdq}\subset r^{dq}=r^q$ (3.22); if $r\subset r_1$ and r_1 is a dq-proximity then $r^q\subset r_1^q=r_1^{dqq}\subset r^{dq}=r_1$ (3.22).

c) $r\subset r^{dq}$; by a) and b), r^{dq} is a dq-proximity; if $r\subset r_1$ and r_1 is a dq-proximity then $r^d\subset r_1$ by a) and 4.8, thus $r^{dq}\subset r_1$ by b). □

4.11 REMARK. If r is a dq-proximity then $r^c=r^{c_0}=r^e\subset r$ is a dq-proximity, but not necessarily a quasi-proximity.

4.12 DEFINITION. If r is a disjointness relation and $S\subset \operatorname{fund} r$ then

$$\mathrm{c}_r^1\, S = \{y\colon\ S\,\bar{r}\,\{y\}\},\quad \mathrm{c}_r^2\, S = \{y\colon\ \{y\}\,\bar{r}\,S\}.$$

4.13 LEMMA. *If r is a quasi-proximity then* $\mathrm{cl}_r^i=\mathrm{c}_r^i$ $(i=1, 2)$.

PROOF. $r=r^{dq}$, thus $A\in\mathscr{T}_r^1$ iff $A\,r\,\{y\}$ for each $y\notin A$.

1° First assume $y\in \mathrm{c}_r^1 S$. Then $S\,\bar{r}\,\{y\}$, thus $\mathrm{cl}_r^1 S\,\bar{r}\,\{y\}$ (by Q1), i.e. $y\in\mathrm{cl}_r^1 S$. Hence $\mathrm{c}_r^1 S\subset \mathrm{cl}_r^1 S$.

2° To prove $\mathrm{cl}_r^1 S\subset \mathrm{c}_r^1 S$, it is enough to show that $\mathrm{c}_r^1 S\in\mathscr{T}_r^1$, i.e. that $\mathrm{c}_r^1 S\,r\,\{y\}$ whenever $y\notin \mathrm{c}_r^1 S$. If $y\notin \mathrm{c}_r^1 S$ then $S\,r\,\{y\}$, so Q3 implies that there are A' and B' with $A'\cup B'=\operatorname{fund} r$, $A' r\,\{y\}$, $S\,r\,B'$. If $z\in B'$ then $S\,r\,\{z\}$ (by Q1), thus $z\notin \mathrm{c}_r^1 S$, hence $\mathrm{c}_r^1 S\subset A'$. Now $\mathrm{c}_r^1 S\,r\,\{y\}$ follows from $A'\,r\,\{y\}$. □

4.14 REMARKS. a) If r is a dq-proximity (d-proximity) then c_r^1 and c_r^2 are closures in the sense of [2] (semi-closures in the sense of [36]).

b) 4.13 states that 3.46 is equivalent to the usual definition of the bitopology induced by r (cf. e.g. [34], [21]) if r is a quasi-proximity. The next lemma is well-known.

4.15 LEMMA. *If r is a quasi-proximity and $A\ r\ B$ then* $\mathrm{cl}_r^1 A\ r\ \mathrm{cl}_r^2 B$ *(i.e. $r=r^{wd}$).*

PROOF. Take A' and B' such that $A'\ r\ B$, $A\ r\ B'$, $A'\cup B'=\mathrm{fund}\ r$. By 4.13, $\mathrm{cl}_r^1 A\cap B'=\emptyset$, so $\mathrm{cl}_r^1 A\subset A'$; hence $\mathrm{cl}_r^1 A\ r\ B$. □

4.16 LEMMA. *If r is a quasi-proximity then* $\mathrm{cl}_{r^w}^i=\mathrm{cl}_r^1$ $(i=1, 2)$.

PROOF. $r=(r^w)^d$ by 4.15, so 3.48 applies. □

§ 5. Generalized preproximities

5.0 LEMMA. *If r is a disjointness relation and $r\subset r^e$ then $r^e=r^{dq}$ is a quasi-proximity; it is the coarsest dq-proximity finer than r.*

PROOF. $r^{ee}\subset r^e$ (since $ee<e$ by 3.36) and $r^e\subset r^{ee}$ (since e is monotone), therefore r^e is a quasi-proximity (4.7). $r^e\subset r^{dq}$ (since $e<dq$); $r^{dq}\subset r^e$ and the second part of the lemma follow from 4.10 c) and 4.8. □

5.1 LEMMA. *If $a=cdq$, c_0dq or dc_0q, and $r\subset r^a$ then $r^a=r^e$. In particular,* 5.0 *remains true if e is replaced by a.*

PROOF. $a<e$ (3.36), so $r^a\subset r^e$. $r^e\subset r^{ae}$ (since e is monotone), thus if $a=\tilde{c}dq$ ($\tilde{c}=c$ or c_0) then $r^e\subset r^{\tilde{c}dqe}\subset r^{\tilde{c}dq}=r^a$ (since $dqe<dq$). According to 3.29, $dc_0q=$ $=c_0dq$. □

5.2 EXAMPLES. a) $r\subset r^e$ does not imply that $r\subset r^a$ where a is as in 5.1: take the relation r from 3.34. (If $A\ r^a\ B$ then $A=\emptyset$ or $B=\emptyset$.)

b) If $r\subset r^a$ with $a=\tilde{c}qd$, $q\tilde{c}d$ or $qd\tilde{c}$ ($\tilde{c}=c_0$ or c) then r^a is not necessarily a dq-proximity: let $|\mathrm{fund}\ r|=4$, and $A\ r\ B$ iff $A\cap B=\emptyset$ and $|A|$ and $|B|$ are even numbers.

5.3 CONVENTIONS. Throughout this section, a fundamental set X will be fixed, and the letters r and s (with or without indices) will denote disjointness relations on X.

For a subset A of X, $y\notin A$ will mean $y\in X\setminus A$.

5.4 DEFINITION. Let a, b_1 and b_2 be operations. r is an *(a, b_1, b_2)-preproximity* if

P3. $r\subset r^a$;

P4. $y\notin A\in \mathrm{dom}\ r\Rightarrow A\ r^{b_1}\ \{y\}$;

P5. $x\notin B\in \mathrm{ran}\ r\Rightarrow \{x\}\ r^{b_2}\ B$.

An *(a, b)-preproximity* is a symmetrical (a, b, b^{-1})-preproximity.

5.5 REMARKS. a) Observe that Axioms P4 and P5 say the same for an (a, b)-preproximity.

b) The preproximities in the sense of Hamburger (cf. 2.0) are the $((c\wedge h)qd, dq)$-preproximities. The strong preproximities introduced in [26] are the $((c\wedge h)qd, d_2)$-preproximities.

5.6 EXAMPLE. A (cqd, dq)-preproximity that is not a $((c\wedge h)qd, dq)$-preproximity: let $X=[0, 1]$, and r the coarsest symmetrical disjointness relation for which

$$[0, x]\, r\, [y, 1] \qquad (0 < x < y < 1),$$
$$([0, x]\cup[y, 1])\, r\, \{z\} \qquad (0 < x < z < y < 1).$$

5.7 LEMMA. *The (e, e, e)-preproximities [(e, e)-preproximities] are the same as the (e, dq, dq)-preproximities [(e, dq)-preproximities].*

PROOF. If $A\, r^e\, \{y\}$ then $A\, r^{dq}\, \{y\}$, as $e<dq$. If $A\, r^{dq}\, \{y\}$ then $A\, r^{edq}\, \{y\}$ by P3, so 3.36 implies $A\, r^e\, \{y\}$. □

5.8 LEMMA. *If r is an (a, b_1, b_2)-preproximity, b_1 and b_2 are elementary operations, then* dom r *is a subbase for $\mathcal{T}_r^1$,* ran r *is a subbase for $\mathcal{T}_r^2$.*

PROOF. If $y\notin A\in\operatorname{dom} r$ then $A\, r^{b_1}\, \{y\}$, thus $A\, r^{dq}\, \{y\}$ by 3.54 a); hence $A\in\mathcal{T}_r^1$. Apply now 3.49. □

5.9 LEMMA. *If b_1 and b_2 are elementary operations and $a<e$ then each (a, b_1, b_2)-preproximity is an (e, e, e)-preproximity.* □

5.10 THEOREM. *Let s be a quasi-proximity and r an (e, e, e)-preproximity.*

a) *$s^w\subset s$ is a (c_0, d_2, d_1)-preproximity and $\mathcal{T}_{s^w}^i=\mathcal{T}_s^i$ $(i=1, 2)$. s^w is the finest (e, e, e)-preproximity coarser than s.*

b) *$r^e=r^{dq}\supset r$ is a quasi-proximity and $\mathcal{T}_{r^e}^i=\mathcal{T}_r^i$ $(i=1, 2)$. r^e is the coarsest dq-proximity finer than r.*

c) *$s^{we}=s$, $r^{ew}\supset r$.*

d) *If s (r) is symmetrical then so is s^w (r^e).*

e) *If $\mathcal{U}_r$ denotes the totally bounded quasi-uniformity*[2] *compatible with r^e then*

$$\{X\times X\setminus A\times B\colon\ A\, r\, B\}$$

is a subbase for $\mathcal{U}_r$.

f) *Let r_1 be another (e, e, e)-preproximity on X. Then $r^e\subset r_1^e$ iff $r\subset r_1^e$.*

PROOF. a) $s^w\subset s$ (3.52 b)), $\mathcal{T}_{s^w}^i=\mathcal{T}_s^i$ (4.16). If $A\, s^w\, B$ then $A\, s\, B$, so there are A' and B' with $A'\cup B'=X$, $A\, s\, B'$, $A'\, s\, B$. By 4.15, $A\, s\, \mathrm{cl}_s^2\, B'$ and $\mathrm{cl}_s^1\, A'\, s\, B$. $A\, s^w\, B$ implies that $A\in\mathcal{T}_s^1$ and $B\in\mathcal{T}_s^2$, therefore $A\, s^w\, \mathrm{cl}_s^2\, B'$ and $\mathrm{cl}_s^1\, A'\, s^w\, B$, i.e. P3 holds for s^w with $a=c_0$. If $y\notin A\in\operatorname{dom} s^w$ then $A\in\mathcal{T}_s^1$, thus $A\, s\, \{y\}$ (4.13), therefore $A\, s\, \mathrm{cl}_s^2\, \{y\}$ (4.15) and $A\, s^w\, \mathrm{cl}_s^2\, \{y\}$, i.e. P4 holds for s^w with $b_1=d_2$. Similarly, P5 holds with $b_2=d_1$, hence s^w is a (c_0, d_1, d_1)-preproximity, and also an (e, e, e)-preproximity (5.9).

Let $r_0\subset s$ be an (e, e, e)-preproximity; we have to show that $r_0\subset s^w$. Assume $A\, r_0\, B$. If $y\notin A$ then $A\, r_0^e\, \{y\}$ by P4, thus $A\, s^e\, \{y\}$ (since e is monotone), i.e. $A\, s\, \{y\}$ (4.7); therefore $A\in\mathcal{T}_s^1$. Similarly, $B\in\mathcal{T}_s^2$. Finally, $r_0\subset s$ implies $A\, s\, B$, thus $A\, s^w\, B$. Hence $r_0\subset s^w$.

---

[2] For definitions and statements concerning (totally bounded) quasi-uniformities and for their connexions with quasi-proximities, see [4], [5], [38], [31] (or think only of the case when r is symmetrical and $\mathcal{U}_r$ is a uniformity).

b) 5.0 and 3.48.

c) $r \subset r^{ew}$ and $s^{we} \subset s$ follow from a) and b). Furthermore, $s = s^{wd} = s^{wde} \subset s^{we}$ (4.15, 4.7, 3.36).

d) 3.34, 3.52 a) and 3.42 b).

e) Put

$$U_{AB} = X \times X \setminus A \times B.$$

$\{U_{AB}\colon A\, r^e\, B\}$ is a subbase for $\mathscr{U}_r$; to prove that $\{U_{AB}\colon A\, r\, B\}$ is a subbase, it is enough to show that if $A\, r^e\, B$ then there are A_i and B_i with $A_i\, r\, B_i$ $(1 \leq i \leq n)$ such that $\bigcap_i U_{A_i B_i} \subset U_{AB}$, or equivalently

$$A \times B \subset \bigcup_i A_i \times B_i. \tag{5.11}$$

By b), $r^e = r^{dq}$, so there is, according to 3.39, a finite collection $\mathscr{F}$ such that for each $x \in A$ and $y \in B$, there are $A', B' \in \mathscr{F}$ with $x \in A'$, $y \in B'$, $A'\, r\, B'$; now if $\{(A_i, B_i)\colon 1 \leq i \leq n\}$ consists of all such pairs (A', B') then $A_i\, r\, B_i$, and (5.11) holds.

f) If $r^e \subset r_1^e$ then $r \subset r^e \subset r_1^e$ (P3). Conversely, if $r \subset r_1^e$ then $r^e \subset r_1^{ee} \subset r_1^e$ (3.36). □

5.12 Definition. If r is an (e, e, e)-preproximity then r^e is the quasi-proximity *induced by* r. In other words, r is *compatible with* r^e.

5.13 Corollary. a) *Two (e, e, e)-preproximities r_1 and r_2 induce the same quasi-proximity iff $r_1 \subset r_2^e$ and $r_2 \subset r_1^e$.*

b) *Given a quasi-proximity s, s^w is the finest (e, e, e)-preproximity compatible with s.*

c) *If an (e, e, e)-preproximity is compatible with a quasi-proximity then they induce the same bitopology.*

Proof. b) s^w is compatible with s, since $s^{we} = s$ (5.10 c)). If r_0 is an (e, e, e)-preproximity compatible with s then $r_0^e = s$, thus $r_0 \subset s$ (P3), therefore $r_0 \subset s^w$ by 5.10 a). □

5.14 Corollary. a) *A bitopology is completely regular*[3] *iff it can be induced by an (e, e, e)-preproximity $[(c_0, d_2, d_1)$-preproximity$]$.*

b) *A topology is completely regular iff it can be induced by an (e, e)-preproximity $[(c_0, d_2)$-preproximity$]$.* □

Our next aim is to show that the notion of an (e, e, e)-preproximity cannot be, in a certain sense, further generalized without losing the validity of this important corollary.

[3] For the definition of complete regularity in bitopological spaces, see e.g. [33], [18], [34], [12]; it is here sufficient to know that a bitopology is completely regular iff it has a compatible quasi-proximity, see e.g. [33], [18], [34], [7], [29], [13].

5.15. THEOREM. *Let a, b_1, b_2 and b be elementary operations.*

a) *If $\mathcal{T}_r^1$ is completely regular with respect to*[4] *$\mathcal{T}_r^2$ for each (a, b_1, b_2)-preproximity r then each (a, b_1, b_2)-preproximity is an (e, e, e)-preproximity.*

b) *If $\mathcal{T}_r$ is completely regular for each (a, b)-preproximity r then each (a, b)-preproximity is an (e, e)-preproximity.*

PROOF. 1° First assume $a<e$. Then a) follows from 5.9. To prove b), take $b_1=b$ and $b_2=b^{-1}$.

2° Assume now $a \not< e$. We are going to prove the theorem by producing a topological space $(X, \mathcal{T})$ which is not completely regular and a symmetrical (a, b_1, b_2)-preproximity r compatible with $\mathcal{T}$. Indeed, $\mathcal{T}_r^1=\mathcal{T}$ is now not completely regular with respect to $\mathcal{T}_r^2=\mathcal{T}$ by Footnote 4, so a) holds; b) follows again from a) just like in 1°.

Consider a regular T_0-space $(X, \mathcal{T})$ which is not completely regular, and let $r=r_{\mathcal{T}}$ (3.44). We claim that r is a symmetrical (a, b_1, b_2)-preproximity compatible with $\mathcal{T}$.

r is a symmetrical disjointness relation with $\operatorname{dom} r=\mathcal{T}$ (3.45). If $y \notin A \in \operatorname{dom} r= =\mathcal{T}$ then $A\, r\, \{y\}$ (since a regular T_0-space is T_1), thus $A\, r^{dq}\, \{y\}$. Hence $A \in \mathcal{T}_r$, i.e. $\operatorname{dom} r \subset \mathcal{T}_r$. Now 3.49 implies that $\operatorname{dom} r=\mathcal{T}$ is a subbase for $\mathcal{T}_r$, i.e. $\mathcal{T}=\mathcal{T}_r$. So r is compatible with $\mathcal{T}$.

3° *Proof of* P3. If $A\, r\, B$ then $A, B \in \mathcal{T}=\mathcal{T}_r$, so $A\, r^w\, B$; consequently,

$$r = r^w. \tag{5.16}$$

By 3.54, there is an n with $w^n<a$, so $r=r^w=r^{w^n}\subset r^a$, i.e. P3 holds.

4° *Proof of* P4 *and* P5. It is enough to show that (with $b=b_1$ or b_2):

$$\exists n>0,\ r^{c_0^n}\subset r^b; \tag{5.17}$$

$$n>0,\ y\notin A\in\mathcal{T} \Rightarrow A\, r^{c_0^n}\, \{y\}. \tag{5.18}$$

(Indeed, if (5.17) and (5.18) hold then $A\, r^b\, \{y\}$ whenever $y \notin A \in \operatorname{dom} r=\mathcal{T}$.)

Instead of (5.17), we shall prove the following stronger statement (the proof by induction does not work for (5.17)):

$$\exists n>0,\ \forall k\geqq 0,\ r^{c_0^{k+n}}\subset r^{c_0^k b}, \tag{5.19}$$

where c_0^0 means h.

5° *Proof of* (5.18). For the sake of simplicity, we shall consider only the case $n=2$. Assume $y \notin A \in \mathcal{T}$, and take closed sets F_i and open sets G_i such that

$$X\setminus A \supset F_1 \supset G_1 \supset F_2 \supset G_2 \supset F_3 \supset G_3 \ni y.$$

$(X\setminus G_3)\ r\ \{y\}$ (since $\mathcal{T}$ is T_1), $(X\setminus G_2)\ r\ F_3$ and $F_3\cup(X\setminus G_3)=X$, therefore $(X\setminus G_2)\, r^{c_0}\, \{y\}$. Similarly, $A\, r^{c_0}\, F_2$, thus $A\, r^{c_0 c_0}\, \{y\}$ follows from $F_2\cup(X\setminus G_2)=X$. (In general, one has to take G_i and F_i for $1\leqq i\leqq 2^n-1$.)

[4] "$\mathcal{P}$ is completely regular with respect to $\mathcal{Q}$" is a statement weaker than "$(\mathcal{P}, \mathcal{Q})$ is completely regular" (see [33], [12]); we shall only use in the proof that a topology is completely regular with respect to itself iff it is completely regular in the usual sense.

6° *Preliminaries to the proof of* (5.19). α) $r^{c_0^n}$ is compatible with $\mathscr{T}$ $(n=1, 2, \ldots)$. Indeed, if $A\, r^{c_0^n}\, B$ then there is a B' with $A\, r^{c_0^{n-1}}\, B'$, therefore

$$\operatorname{dom} r^{c_0^n} \subset \operatorname{dom} r^{c_0^{n-1}} \subset \ldots \subset \operatorname{dom} r = \mathscr{T};$$

on the other hand, (5.18) implies $\mathscr{T} \subset \operatorname{dom} r^{c_0^n}$, i.e.

$$\operatorname{dom} r^{c_0^n} = \mathscr{T}. \tag{5.20}$$

By (5.18), $A\,(r^{c_0^n})^{dq}\,\{y\}$ whenever $y \notin A \in \operatorname{dom} r^{c_0^n} = \mathscr{T}$, hence $r^{c_0^n}$ is compatible with $\mathscr{T}$ (3.49).

β) For $k=1, 2, \ldots,$ we have

$$r^{c_0^k} = r^{c_0^k w}. \tag{5.21}$$

Indeed, if $A\, r^{c_0^k}\, B$ then $A, B \in \mathscr{T}$ by (5.20), thus, according to α), A and B are closed in the topology induced by $r^{c_0^k}$, therefore $A\, r^{c_0^k w}\, B$.

γ) If $A\, r^{c_0}\, B$ then $A \cap B = \emptyset$ and, by (5.20), $A, B \in \mathscr{T}$, thus $A\, r\, B$; hence

$$r^{c_0} \subset r. \tag{5.22}$$

7° *Proof of* (5.19). α) If $b=h, d_1, d_2, d, q_1, q_2, q, c_0, c$ or w then (5.19) holds with $n=j$, i.e.

$$r^{c_0^{k+1}} \subset r^{c_0^k b} \quad (k = 0, 1, \ldots). \tag{5.23}$$

This is evident for $b=c_0$, and follows from 3.26 b) for $b=c$. If $b=h, d_1, d_2, d, q_1, q_2, q$ or w then $w<b$ (3.52 b), 3.14 c), 3.16 c), 3.18 b)), so it is enough to prove

$$r^{c_0^{k+1}} \subset r^{c_0^k w} \quad (k = 0, 1, \ldots) \tag{5.24}$$

instead of (5.23). For $k=0$, this follows from (5.22) and (5.16). If $k \neq 0$ then (5.21) and 3.26 d) imply (5.24).

β) Assume now that (5.19) holds for a_1 and a_2, i.e. there are $i>0$ and $j>0$ such that

$$r^{c_0^{k+i}} \subset r^{c_0^k a_1}, \quad r^{c_0^{k+j}} \subset r^{c_0^k a_2} \quad (k = 0, 1, \ldots). \tag{5.25}$$

β₁) Let $b=a_1 a_2$. Then

$$r^{c_0^{k+i+j}} \underset{1}{\subset} r^{c_0^{k+i} a_2} \underset{2}{\subset} r^{c_0^k a_1 a_2}$$

[1: the second part of (5.25); 2: the first part of (5.25), using that a_2 is monotone]; thus (5.19) holds with $n=i+j$.

β₂) If $b=a_1 \vee a_2$ then

$$r^{c_0^k(a_1 \vee a_2)} = r^{c_0^k a_1} \cup r^{c_0^k a_2} \supset r^{c_0^k a_1} \supset r^{c_0^{k+i}},$$

i.e. (5.19) holds now with $n=i$.

β₃) If $b=a_1 \wedge a_2$ then

$$r^{c_0^k(a_1 \wedge a_2)} = r^{c_0^k a_1} \cap r^{c_0^k a_2} \supset r^{c_0^{k+i}} \cap r^{c_0^{k+j}} \supset r^{c_0^{k+n}}$$

where $n=\max\{i, j\}$, and the last inclusion follows from 3.26 d). □

If not only Corollary 5.14, but also the gist of Theorem 5.10 is required to be true then the notion of an (e, e, e)-preproximity is the most general one, without reservations:

5.26 THEOREM. *The disjointness relation r is an (e, e, e)-preproximity iff*
(i) r^{dq} *is a quasi-proximity;*
(ii) dom r *is a subbase for* $\mathcal{T}_r^1 = \mathcal{T}_{r^{dq}}^1$;
(iii) ran r *is a subbase for* $\mathcal{T}_r^2 = \mathcal{T}_{r^{dq}}^2$.
Condition (ii) *can be replaced by* P4 *with* $b_1 = dq$:
(ii*) dom $r \subset \mathcal{T}_r^1 (= \mathcal{T}_{r^{dq}}^1)$.
Condition (iii) *can be replaced by* P5 *with* $b_2 = dq$:
(iii*) ran $r \subset \mathcal{T}_r^2 (= \mathcal{T}_{r^{dq}}^2)$.

PROOF. (ii) and (ii*), respectively (iii) and (iii*) are equivalent by 3.49.

1° If r is an (e, e, e)-preproximity then (i) follows from 5.10 b), (ii*) and (iii*) from 5.7.

2° Assume now that r satisfies (i), (ii) and (iii).

$$r \underset{1}{\subset} r^{dq} \underset{2}{\subset} r^{dqc_0} \underset{3}{=\!=} r^e$$

[1: evident; 2: (i) and 4.1 Q3; 3: 3.32], thus r satisfies P3 with $a = e$. 5.7 implies that r is an (e, e, e)-preproximity. □

5.27 REMARKS. a) r is an (e, e, e)-preproximity iff $r \subset r^{ew}$.

b) The above proof shows that r^{dq} is *a* quasi-proximity iff $r \subset r^e$. Starting from this observation, one can generalize the notion of a *proximity subbase* in the sense of Sharma [43].

5.28 COROLLARY. *A symmetrical disjointness relation r is an (e, e)-preproximity iff $r^{dq} = s$ is a proximity and* dom $r \subset \mathcal{T}_s$ *(equivalently:* dom r *is a subbase for* $\mathcal{T}_s$). □

5.29 EXAMPLE. In 5.26 and 5.28, dq cannot be replaced by e. Let $X = \mathbf{R}$, and s the proximity induced by the Euclidean metric. Pick disjoint closed sets A_0 and B_0 such that $A_0 \bar{s} B_0$. Put

$$r = s^w \cup \{(A_0, B_0), (B_0, A_0)\}.$$

Now $r^e = s$, dom $r =$ ran r is a subbase for $\mathcal{T}_r = \mathcal{T}_{r^{dq}} = \mathcal{T}_s$, but r is not an (e, e, e)-preproximity, since $A_0 \, r \, B_0$ and $A_0 \, \overline{r^e} \, B_0$.

5.30 EXAMPLE. If $\{a_1, a_2, a_3\} = \{\bar{c}, d, q\}$ ($\bar{c} = c_0$ or c) and $a_1 a_2 a_3 \neq e$ then the notion of an (e, e)-preproximity is strictly more general than that of an $(a_1 a_2 a_3, dq)$-preproximity. This is shown by the relation r from 3.34. A less simple but more interesting example will be given in 5.37.

5.31 THEOREM. *If r is an (e, e, e)-preproximity* [*(e, e)-preproximity*] *then so is r^q. r and r^q induce the same quasi-proximity* [*proximity*].

PROOF. 5.26 and $qdq = dq$. (5.26 (ii*) holds for r^q because each element of dom r^q is a finite union of elements of dom r.) □

5.32 Remarks. a) A similar statement holds for (cdq, dq, dq)-preproximities. (If $r \subset r^{cdq}$ then $r^q \subset r^{cdqq} = r^{cdq} \subset (r^q)^{cdq}$.)

b) Observe that $r = r^q$ for the (e, e)-preproximity r in 5.30.

c) The same example shows that for an (e, e)-preproximity r, even $\operatorname{dom} r^q$ is not necessarily a base for $\mathcal{T}_r = \mathcal{T}_{r^q}$.

5.33 Proof of Theorem 1.9. $r_{\mathcal{S}}$ is a symmetrical disjointness relation and $\operatorname{dom} r$ is a subbase for $\mathcal{T}$ (3.45). $r_{\mathcal{S}}$ is, according to the conditions of the theorem, an (e, dq)-preproximity, so it is an (e, e)-preproximity (5.7). $r_{\mathcal{S}}$ is compatible with $\mathcal{T}$ (5.8), so $\mathcal{T}$ is completely regular (5.14b)). □

5.34 Remarks. a) Several equivalent versions of Theorem 1.9 can be formulated, using Lemmas 3.32, 3.38, 3.39 and 3.40.

b) If the subbase $\mathcal{S}$ satisfies the conditions of Theorem 1.9 [1.6] then so does the base generated by $\mathcal{S}$; this can be deduced from 5.31 [5.32a)], but a direct proof is simpler. Therefore if these theorems are known to be true for bases then they can be easily proved for subbases, too. The analogous statement for Theorem 1.1 is false, see Example 5.35 below.

c) A bitopological equivalent of Theorem 1.9 can be obtained considering the relation r defined by

$$A\, r\, B \Leftrightarrow A \in \mathcal{S}_1^*,\ B \in \mathcal{S}_2^*,\ A \cap B = \emptyset$$

where $\mathcal{S}_i^* = \mathcal{S}_i \cup \{\emptyset, X\}$ and $\mathcal{S}_1$ and $\mathcal{S}_2$ are subbases for the two topologies. In this way one can generalize results from [41], [14], [42], [15], [13], [51].

It is left to the reader to state and prove the theorem.

5.35 Example. Let $X = \{x_i, y_i, z_i \colon i = 1, 2, 3, 4\}$,

$$\mathcal{S} = \big\{\{x_i, y_{i+1}, z_i\}, \{x_{i+1}, y_i, z_i\} \colon 1 \leqq i \leqq 4\big\} \cup$$
$$\cup \big\{\{x_1, x_2, x_3, x_4\}, \{y_1, y_2\}, \{y_3, y_4\}\big\}$$

where $x_5 = x_1$ and $y_5 = y_1$. $\mathcal{S}$ is a subbase for the discrete topology on X, and it satisfies the conditions of Theorem 1.1. The base generated by $\mathcal{S}$, however, does not satisfy 1.1 (iii) (consider the sets $\{x_1, x_2, x_3, x_4\}$ and $\{y_1, y_2, y_3, y_4\}$).

This example also shows that 5.31 is not valid for (cqd, dq, dq)-preproximities or $((c \wedge h)qd, dq, dq)$-preproximities. (Observe that, in contrast to 5.6, a relation of the form $r_{\mathcal{S}}$ is a $((c \wedge h)qd, dq, dq)$-preproximity iff it is a (cqd, dq, dq)-preproximity.)

5.36 Lemma. *If $\mathcal{S}$ is a family of subsets of X and $r_{\mathcal{S}, X} = r_{\mathcal{S}} \subset r_{\mathcal{S}}^{qdc}$ then $r_{\mathcal{S}} \subset r_{\mathcal{S}}^{c}$.*

Proof. Assume $A\, r_{\mathcal{S}}\, B$. Then $A\, r_{\mathcal{S}}^{qdc}\, B$, i.e. there is a finite covering $\mathcal{C}$ of X such that for each $C \in \mathcal{C}$, either $A\, r_{\mathcal{S}}^{qd}\, C$ or $B\, r_{\mathcal{S}}^{qd}\, C$. In the first case, there are finite collections $\mathcal{E}_C$ and $\mathcal{F}_C$ such that $A \subset \bigcup \mathcal{E}_C$, $C \subset \bigcup \mathcal{F}_C$, and $E\, r_{\mathcal{S}}\, F$ whenever $E \in \mathcal{S}_C$ and $F \in \mathcal{F}_C$. $E\, r_{\mathcal{T}}\, F$ implies $F \in \mathcal{S}^* = \mathcal{S} \cup \{\emptyset, X\}$ and $E \cap F = \emptyset$, so $A \cap F = \emptyset$; from $A\, r_{\mathcal{S}}\, B$ we have $A \in \mathcal{S}^*$, too, therefore $A\, r_{\mathcal{S}}\, F$. Similarly, if $A\, r_{\mathcal{S}}^{qd}\, C$ then there is a finite collection $\mathcal{F}_C$ such that $C \subset \bigcup \mathcal{F}_C$, and $B\, r_{\mathcal{S}}\, F$ for each $F \in \mathcal{F}_C$. Now the covering $\bigcup_{C \in \mathcal{C}} \mathcal{F}_C$ shows that $A\, r_{\mathcal{S}}^{c}\, B$. □

5.37 Example. There is an (e, e)-preproximity of the form $r_{\mathcal{S}}$ with the same properties as r in 5.30. Take X and $\mathcal{S}$ from 1.11. As shown there, $r_{\mathcal{S}}$ is an (e, e)-preproximity, but

$$r_{\mathcal{S}} \not\subset r_{\mathcal{S}}^{cdq}, \tag{5.38}$$

so $r_{\mathcal{S}}$ is not a (cdq, dq)-preproximity. (5.38) implies $r_{\mathcal{S}} \not\subset r_{\mathcal{S}}^{c}$ (since $c < cdq$), $r_{\mathcal{S}} \not\subset r_{\mathcal{S}}^{qdc}$ (5.36), therefore $r_{\mathcal{S}}$ is not a (qdc, dq)-preproximity. According to the inequalities in 3.34, $r_{\mathcal{S}}$ is not an $(a_1 a_2 a_3, dq)$-preproximity either.

§ 6. The category of (e, e, e)-preproximities

6.0 a) Let C be a concrete category over Set, and denote the underlying functor by $\text{fund} = \text{fund}_{\mathrm{C}}$. An object $\mathfrak{x}$ of C will be called a structure on the set fund $\mathfrak{x}$. Instead of "m: $\mathfrak{x} \to \mathfrak{y}$ is a morphism (in C)", we shall say, as usual, that "fund m is $(\mathfrak{x}, \mathfrak{y})$-continuous (in C)".

For structures $\mathfrak{x}$ and $\mathfrak{y}$ on the same set X, $\mathfrak{x}$ is *finer* than $\mathfrak{y}$ or $\mathfrak{y}$ is *coarser* than $\mathfrak{x}$ ($\mathfrak{y} < \mathfrak{x}$)[5] if id_X is $(\mathfrak{x}, \mathfrak{y})$-continuous.

b) The concrete category C is a *topological category*[6] if

(i) the structures on a set form a set, not a proper class;

(ii) there is at most one structure on a one-point set;

(iii) if $\mathfrak{x} < \mathfrak{y}$ and $\mathfrak{y} < \mathfrak{x}$ then $\mathfrak{x} = \mathfrak{y}$;

(iv) if $\mathfrak{X}$ is a (possibly empty) family of structures on X then there is a coarsest structure $\sup \mathfrak{X} = \sup_{\mathrm{C}} \mathfrak{X} = \sup_{\mathrm{C},X} \mathfrak{X}$ on X finer than each element of $\mathfrak{X}$;

(v) if f: $X \to Y$ and $\mathfrak{y}$ is a structure on Y then there is a structure $f^{-1}\mathfrak{y} = f_{\mathrm{C}}^{-1}\mathfrak{y}$ on X such that f is $(\mathfrak{x}, \mathfrak{y})$-continuous iff $f^{-1}\mathfrak{y} < \mathfrak{x}$;

(vi) if f: $X \to Y$ and $\mathfrak{Y}$ is a family of structures on Y then $f^{-1} \sup \mathfrak{Y} = \sup \{f^{-1}\mathfrak{y} : \mathfrak{y} \in \mathfrak{Y}\}$;

(vii) if f: $X \to Y$, g: $Y \to Z$ and $\mathfrak{z}$ is a structure on Z then $f^{-1}g^{-1}\mathfrak{z} = (g \circ f)^{-1}\mathfrak{z}$.

c) (vi) can be replaced by: (vi*) if f: $X \to Y$ and $\mathfrak{x}$ is a structure on X then there is a structure $f\mathfrak{x} = f_{\mathrm{C}}\mathfrak{x}$ on Y such that f is $(\mathfrak{x}, \mathfrak{y})$-continuous iff $\mathfrak{y} < f\mathfrak{x}$.

d) We list below some good properties of topological categories; see [28] and [29] for more.

1° $\sup \mathfrak{X}$ in (iv), $f^{-1}\mathfrak{y}$ in (v) and $f\mathfrak{x}$ in (vi*) are uniquely determined.

2° The structures on a set form a complete lattice, i.e. inf can be defined analogously to sup.

3° On each set, there exist a finest and a coarsest structure (called *discrete*, respectively *indiscrete*).

4° Restrictions of structures to subsets can be defined as a special case of (v).

---

[5] Several authors write $\mathfrak{x} \leqq \mathfrak{y}$ instead of $\mathfrak{y} < \mathfrak{x}$.

[6] There are a number of equivalent versions of this definition; the one given here is essentially from [9]. For other possible definitions, see the references in [29] 1.1. Topological categories are sometimes called "properly fibred topological categories", e.g. in [28] and [9].

5° Arbitrary products and sums do exist in C.
6° The constant mappings are continuous.

e) Let C and D be concrete categories. A *concrete functor* a: C → D can be interpreted as a function assigning to each C-structure $\mathfrak{x}$ a D-structure $\mathfrak{x}^a$ on the same set such that if f is $(\mathfrak{x}, \mathfrak{y})$-continuous in C then it is also $(\mathfrak{x}^a, \mathfrak{y}^a)$-continuous in D. (To make a a functor in the usual sense, define it on the C-morphisms by $\mathrm{fund}_{\mathsf{D}}\, m^a = \mathrm{fund}_{\mathsf{C}}\, m$.)

f) *Subcategory* will always mean a full, isomorphism-closed subcategory. From now on, let C always denote a topological category, and let D be a subcategory of C.

g) D is sup-*closed*, respectively inf-*closed* (in C) if for an arbitrary family $\mathfrak{X}$ of D-structures on the same set, $\sup_{\mathsf{C}} \mathfrak{X}$, respectively $\inf_{\mathsf{C}} \mathfrak{X}$ is a D-structure.

h_1) D is sup-closed iff for each C-structure $\mathfrak{x}$ on X, there is a finest D-structure $\mathfrak{x}^v$ coarser than $\mathfrak{x}$.

h_2) D is inf-closed iff for each C-structure $\mathfrak{x}$ on X, there is a coarsest D-structure $\mathfrak{x}^u$ finer than $\mathfrak{x}$.

i_1) If D is sup-closed in C then $\sup_{\mathsf{D}}$, $\inf_{\mathsf{D}}$ and f_{D} exist, and

$$\sup_{\mathsf{D}} \mathfrak{X} = \sup_{\mathsf{C}} \mathfrak{X}, \quad \inf_{\mathsf{D}} \mathfrak{X} = (\inf_{\mathsf{C}} \mathfrak{X})^v, \quad f_{\mathsf{D}}\, \mathfrak{x} = (f_{\mathsf{C}}\, \mathfrak{x})^v.$$

i_2) If D is inf-closed then

$$\inf_{\mathsf{D}} \mathfrak{X} = \inf_{\mathsf{C}} \mathfrak{X}, \quad \sup_{\mathsf{D}} \mathfrak{X} = (\sup_{\mathsf{C}} \mathfrak{X})^u, \quad f_{\mathsf{D}}^{-1}\, \mathfrak{y} = (f_{\mathsf{C}}^{-1}\, \mathfrak{y})^u.$$

j_1) If D is inf-closed and $\{\mathfrak{x}_i \colon i \in I\}$ is a family of C-structures then

$$\sup_{\mathsf{D}\, i} \mathfrak{x}_i^u = (\sup_{\mathsf{C}\, i} \mathfrak{x}_i)^u.$$

k_1) D is *inverse-closed* if $f_{\mathsf{C}}^{-1}\mathfrak{y}$ is a D-structure whenever $\mathfrak{y}$ is a D-structure.

l_1) D is inverse-closed iff f_{D}^{-1} exists and $f_{\mathsf{D}}^{-1}\mathfrak{y} = f_{\mathsf{C}}^{-1}\mathfrak{y}$.

j_2) to l_2) A dual definition of *image-closed* subcategories, respectively dual statements.

m) D is *bireflective* (in C) if it is sup-closed and v is a concrete functor; v is called the *reflector* (for D). D is *bicoreflective* if it is inf-closed and u is a concrete functor; u is the *coreflector*.

n) D is bireflective iff it is sup-closed and inverse-closed; D is bicoreflective iff it is inf-closed and image-closed.

o) D is bicoreflective iff it is inf-closed, and one of the following equivalent conditions holds in C:

(1) $$(f^{-1}\, \mathfrak{y})^u = (f^{-1}\, \mathfrak{y}^u)^u;$$

(2) $$(f^{-1}\, \mathfrak{y})^u > f^{-1}\, \mathfrak{y}^u.$$

p) A bi(co)reflective subcategory of a topological category is topological.

q) D will be called *I-bicoreflective* if it is bicoreflective and inverse-closed.

r) D is I-bicoreflective iff it is inf-closed, and the following holds in C:

$$(f^{-1}\mathfrak{x})^u = f^{-1}\mathfrak{x}^u.$$

s) D will be called *S-bicoreflective* if it is bicoreflective, and closed under taking substructures as well as under taking suprema of non-empty collections.

t) If a is a concrete functor from C into C′ then the subcategory of C′ whose structures are of the form $\mathfrak{x}^a$ ($\mathfrak{x}$ a C-structure) will be denoted by C^a. $a|D$ is the restriction of a to the subcategory D. If a is a concrete functor then so is $a|D$. We shall write D^a instead of $D^{a|D}$.

6.1 Definition and Conventions. Let Dis be the category of disjointness relations, containing $(\emptyset, \emptyset)$, such that f: fund $r \to$ fund s is (r, s)-continuous iff $P\,s\,Q$ implies $f^{-1}[P]\,r\,f^{-1}[Q]$.

From now on, operations will be regarded as restricted to Dis. a-proximities and (a, b_1, b_2)-preproximities will be required to contain $(\emptyset, \emptyset)$.

6.2 Lemma. Dis *is a topological category,*

$$\sup_{i\in I} r_i = \bigcup_{i\in I} r_i, \qquad \inf_{i\in I} r_i = \bigcap_{i\in I} r_i \quad (I \neq \emptyset),$$

the discrete, respectively the indiscrete relation on X is

$$\delta_X = \{(A, B)\colon A, B \subset X, A \cap B = \emptyset\}, \quad \iota_X = \{(\emptyset, X), (X, \emptyset), (\emptyset, \emptyset)\};$$

$$A\,(f^{-1}s)\,B \Leftrightarrow \exists P, Q, \quad A = f^{-1}[P], \quad B = f^{-1}[Q],\ P\,s\,Q;$$

$$P\,(fr)\,Q \Leftrightarrow P \cap Q = \emptyset, \quad f^{-1}[P]\,r\,f^{-1}[Q]. \ \square$$

6.3 Lemma. *Any elementary operation is a concrete functor from* Dis *into* Dis.

Proof. 1° Straightforward for $h, d, d_1, d_2, q, q_1, q_2, c_0$ and c.
2° Let $f\colon X \to Y$ be (r, s)-continuous. To prove that w is also a concrete functor, we have to show that f is (r^w, s^w)-continuous, too. Assume $P\,s^w\,Q$; then $P\,s\,Q$, so $f^{-1}[P]\,r\,f^{-1}[Q]$. To obtain $f^{-1}[P]\,r^w\,f^{-1}[Q]$, it is enough to prove that $f^{-1}[P] \in \mathcal{T}_r^1$ and $f^{-1}[Q] \in \mathcal{T}_r^2$.
Take a point $z \in X \setminus f^{-1}[P]$. $P\,s^w\,Q$ implies that $P \in \mathcal{T}_s^1$, therefore $P\,s^{dq}\,\{f(z)\}$. By 1° (applied first to d, then to q), $f^{-1}[P]\,r^{dq}\,f^{-1}[\{f(z)\}]$, hence $f^{-1}[P]\,r^{dqd}\,\{z\}$. Now $f^{-1}[P] \in \mathcal{T}_r^1$ follows from $dqd < dq$.
3° An easy induction completes the proof. $\square$

6.4 Consider the following subcategories of Dis:

Dis_σ = the symmetrical disjointness relations,
$a\text{Prox}$ = the a-proximities (a an operation),
PrProx = the (e, e, e)-preproximities,
$a\text{Prox}_\sigma$ = the symmetrical a-proximities,
PrProx_σ = the (e, e)-preproximities,
QProx = eProx (= the quasi-proximities),
Prox = $e\text{Prox}_\sigma$ (= the proximities).

The next theorem sums up well-known results:

6.5 THEOREM. a) dProx *is S-bicoreflective in* Dis, *with the coreflector d. dq*Prox *is I-bicoreflective in d*Prox, *with the coreflector* $q|d$Prox. QProx *is bireflective in dq*Prox.

$$\sup_{i\in I}{}_{d\mathrm{Prox}}\, r_i = \bigcup_{i\in I} r_i, \quad \sup_{i\in I}{}_{dq\mathrm{Prox}}\, r_i = \sup_{i\in I}{}_{\mathrm{QProx}}\, r_i = \Big(\bigcup_{i\in I} r_i\Big)^q \quad (I \neq \emptyset).$$

δ_X *is the discrete,* ι_X^d *the indiscrete relation on X in d*Prox, *dq*Prox *and* QProx.

$$f^{-1}_{\mathrm{QProx}}\, s = f^{-1}_{dq\mathrm{Prox}}\, s = f^{-1}_{d\mathrm{Prox}}\, s = (f^{-1}_{\mathrm{Dis}}\, s)^d.$$

b) *If* C *is any of the above categories then* C_σ *is bireflective and bicoreflective in* C, *with the coreflector* $h\vee(-1)|\mathsf{C}$. *All the statements in* a) *remain valid if each category* C *is replaced by* C_σ. □

6.6 THEOREM. PrProx *is bireflective in* Dis, PrProx_σ *in* Dis_σ. PrProx_σ *is bireflective and bicoreflective in* PrProx. *In particular,* PrProx *and* PrProx_σ *are topological categories, and*

$$\sup_{i\in I}{}_{\mathrm{PrProx}_\sigma}\, r_i = \sup_{i\in I}{}_{\mathrm{PrProx}}\, r_i = \bigcup_{i\in I} r_i \qquad (I \neq \emptyset),$$

$$f^{-1}_{\mathrm{PrProx}_\sigma}\, s = f^{-1}_{\mathrm{PrProx}}\, s = f^{-1}_{\mathrm{Dis}}\, s.$$

The discrete and the indiscrete relations are the same as in Dis. □

6.7 The restriction $r|S$ of a disjointness relation r to a subset S of fund r can be defined, according to 6.2, as follows:

$$r|S = \{(A\cap S, B\cap S)\colon\ A\ r\ B\}.$$

By 6.6, the same holds in PrProx and PrProx_σ, too.

Let now r_i be a disjointness relation on X_i $(i\in I\neq\emptyset)$. The product r of the relations r_i in Dis is defined by

$$r = \bigcup_{i\in I} \big\{(\mathrm{pr}_i^{-1}[A], \mathrm{pr}_i^{-1}[B])\colon\ i\in I,\ A\ r_i\ B\big\},$$

where pr_i is the canonical projection of $\bigtimes_{j\in I} X_j$ onto X_i. By 6.6, the same construction yields the product in PrProx and PrProx_σ.

Compare these statements on restriction and product with [23], Theorems 2.5 and 2.6.

6.8 THEOREM. a) QProx *is isomorphic to an I-bicoreflective subcategory of* PrProx. *More precisely,* $w|$QProx *is a concrete isomorphism; its inverse is* $e|\mathrm{QProx}^w = dq|\mathrm{QProx}^w$; QProx^w *is I-bicoreflective in* PrProx, *with the coreflector* $ew|\mathrm{PrProx} = dqw|\mathrm{PrProx}$.

b) *An analogous statement for* Prox *and* PrProx_σ.

PROOF. 1° By 5.10, w and e (dq) establish a one-to-one correspondence between the quasi-proximities and certain (e, e, e)-preproximities on the same set. According to 6.3, $w|$QProx and its inverse are both concrete functors, thus they are isomorphisms.

2° If r is an (e, e, e)-preproximity then $r^{ew}=r^{dqw}$ is also an (e, e, e)-preproximity, and $r\subset r^{ew}$ (5.10). If r_1 is another (e, e, e)-preproximity on fund r with $r\subset r_1^{ew}$ then $r^{ew}\subset r_1^{ewew}=r_1^{ew}$ (5.10 c)), thus PrProx^{ew} is inf-closed in PrProx (6.0 j$_1$) with $u=$ $=ew|\mathsf{PrProx}$). According to 5.10, $\mathsf{PrProx}^e=\mathsf{PrProx}^{dq}=\mathsf{QProx}$, thus $\mathsf{PrProx}^{ew}=$ $=\mathsf{PrProx}^{dqw}=\mathsf{QProx}^w$.

If s is a disjointness relation then

$$(6.9) \qquad (f_{\mathsf{Dis}}^{-1}\, s)^{dq} = (f_{\mathsf{Dis}}^{-1}\, s^{dq})^d$$

(check directly from the definitions, or deduce from 6.5 a)). If s is an (e, e, e)-preproximity then

$$(f_{\mathsf{PrProx}}^{-1}\, s)^{dq} = f_{\mathsf{QProx}}^{-1}\, s^{dq}$$

((6.9), 6.6 and 6.5). Therefore

$$(f_{\mathsf{PrProx}}^{-1}\, s)^{dqw} = (f_{\mathsf{QProx}}^{-1}\, s^{dq})^w = f_{\mathsf{QProx}^w}^{-1}\, s^{dqw}$$

(the second equality follows from 1°), i.e. QProx^w is I-bicoreflective (6.0 r)). □

6.10 From 6.0 i$_1$), 6.0 l$_2$) and 6.8 we have

$$\Big(\prod_{i\in I}{}_{\mathsf{PrProx}}\, r_i\Big)^e = \prod_{i\in I}{}_{\mathsf{QProx}}\, r_i^e$$

(compare with [23], Theorems 2.6 and 2.7).

6.11 Definition. A disjointness relation r is an *(a, b_1, b_2)-metaproximity* if there is an (a, b_1, b_2)-preproximity s with $s^d=r$. The category of the (e, e, e)-metaproximities will be denoted by MeProx.

6.12 Lemma. a) *For a disjointness relation r, the following are equivalent:*
(i) *r is an (e, e, e)-metaproximity;*
(ii) *r is a wd-proximity and r^q is a quasi-proximity;*
(iii) *r is a wd-proximity and r^w is an (e, e, e)-preproximity.*
b) *If s is an (e, e, e)-preproximity then $s\subset s^{dw}$.*
c) *Each quasi-proximity is an (e, e, e)-metaproximity.*

Proof. a) (i)⇒(ii): Assume that $r=s^d$ where s is an (e, e, e)-preproximity. $r^{wd}=s^{dwd}\subset s^{dd}=s^d=r$ (3.16 b), 3.52 b)). If $A\, r\, B$ then $A'\, s\, B'$ with some $A'\supset A$ and $B'\supset B$. $A'\in\mathscr{T}_s^1$, $B'\in\mathscr{T}_s^2$ (5.8), hence $A'\in\mathscr{T}_r^1$, $B'\in\mathscr{T}_r^2$ (3.48). Moreover, $A'\, s\, B'$ implies $A'\, r\, B'$, thus $A'\, r^w\, B'$ and $A\, r^{wd}\, B$. Therefore $r\subset r^{wd}$, too, i.e. r is a wd-proximity. $r^q=s^{dq}$ is a quasi-proximity by 5.10.

(ii)⇒(iii): We are going to show that r^w satisfies 5.26 (i), (ii)* and (iii)*.

$$(6.13) \qquad (r^w)^{dq} = r^{wdq} = r^q,$$

thus $(r^w)^{dq}$ is a quasi-proximity, i.e. 5.26 (i) holds. If $A\in\operatorname{dom} r^w$ then

$$A\in\mathscr{T}_r^1 \underset{1}{=} \mathscr{T}_{r^q}^1 \underset{2}{=} \mathscr{T}_{(r^w)^{dq}}^1$$

[1: 3.48; 2: (6.13)], i.e. 5.26 (ii)* holds, too.

(iii)⇒(i): Evident.

b) 3.48 and 5.8.

c) 4.15 and a). □

6.14 Remarks. a) If r is an (e, e, e)-metaproximity then r^w is not only an (e, e, e)-preproximity, but also a (qc_0, qd_2, qd_1)-preproximity. This means that, in contrast to the case of the preproximities, replacing $a=cqd$ by $a=cdq$ or even $a=dcq$ is not a proper generalization in the definition of metaproximities.

b) 6.12 c) is also true in a stronger form: each quasi-proximity is a (c_0, d_2, d_1)-metaproximity.

6.15 There is an obvious factorization of the functor $dq|\mathsf{PrProx}$ through MeProx. The following can be proved:

a) MeProx is isomorphic to an I-bicoreflective subcategory of PrProx. More precisely, $w|\mathsf{MeProx}$ is a concrete isomorphism, its inverse is $d|\mathsf{MeProx}^w$, MeProx^w is I-bicoreflective in PrProx, with the coreflector $dw|\mathsf{PrProx}$.

b) QProx is I-bicoreflective in MeProx, with the coreflector $q|\mathsf{MeProx}$.

c) MeProx is bireflective in dProx. Consequently, MeProx is a topological category,

$$\sup_{i\in I}{}_{\mathsf{MeProx}}\, r_i = \bigcup_{i\in I} r_i \quad (I \neq \emptyset),$$

$$f^{-1}_{\mathsf{MeProx}}\, s = (f^{-1}_{\mathsf{Dis}}\, s)^d,$$

δ_X is the discrete, ι_X^d the indiscrete relation on X in MeProx.

d) Products in MeProx can be easily described:

$$A\Big(\prod_{i\in I} r_i\Big)\, B \Leftrightarrow \exists i\in I, \quad \mathrm{pr}_i[A]\, r_i\, \mathrm{pr}_i[B].$$

e) Similarly to 6.10,

$$\Big(\prod_{i\in I}{}_{\mathsf{PrProx}} r_i\Big)^d = \prod_{i\in I}{}_{\mathsf{MeProx}} r_d^i, \quad \Big(\prod_{i\in I}{}_{\mathsf{MeProx}} r_i\Big)^q = \prod_{i\in I}{}_{\mathsf{QProx}} r_i^q.$$

§ 7. Preproximities for Leader and Lodato proximities

7.0 Definition. A dq-proximity r is a *Leader proximity* if[7]

L. $\quad A\, r\, B \Rightarrow \mathrm{c}_r^1 A\, r\, \mathrm{c}_r^2 B.$

A symmetrical Leader proximity is called a *Lodato proximity*. (See [35], [37], [39].)

7.1 Lemma. *For a disjointness relation r, the following conditions are equivalent:*
(i) *r is a Leader proximity;*
(ii) *r is a dq-proximity satisfying*

L*. $\quad A\, r\, B \Rightarrow \mathrm{cl}_r^1 A\, r\, \mathrm{cl}_r^2 B;$

(iii) *r is a dq-proximity and a wd-proximity;*
(iv) *r is a wdq-proximity.*

[7] In the more usual terminology, only $A\, r\, \mathrm{c}_r^2 B$ is required in the definition of a Leader proximity, see [39].

PROOF. (i)$\Rightarrow$(ii): It is well-known that $c_r^i = cl_r^i$ for Leader proximities. (Indeed, if $y \notin cl_r^1 A$ then $cl_r^1 A \, r \, \{y\}$ [since $r = r^{dq}$], so $A \, r \, \{y\}$, i.e. $y \notin c_r^1 A$; if $z \notin c_r^1 A$ then $A \, r \, \{z\}$ and, by L, $c_r^1 A \, r \, \{z\}$, i.e. $c_r^1 A \in \mathcal{T}_r^1$, hence $y \in cl_r^1 A$ implies $y \in c_r^1 A$.)

(ii)$\Rightarrow$(i): The first part of the argument in (i)$\Rightarrow$(ii) gives $c_r^i A \subset cl_r^i A$.

(ii)$\Rightarrow$(iii): Evident.

(iii)$\Rightarrow$(iv): $r^{wdq} = (r^{wd})^q = r^q = r$.

(iv)$\Rightarrow$(ii): $r^{dq} = r^{wdqdq} \subset r^{wdq} = r$, thus r is a dq-proximity. To prove L*, assume $A \, r \, B$. Then $A \, r^{wdq} \, B$, so there are finite collections $\mathcal{A}$ and $\mathcal{B}$ such that $A = \bigcup \mathcal{A}$, $B = \bigcup \mathcal{B}$, and for each pair $A' \in \mathcal{A}$, $B' \in \mathcal{B}$, there are sets $A'' \supset A'$ and $B'' \supset B'$ with $A'' r^w B''$, i.e. $A'' \, r \, B''$, $A'' \in \mathcal{T}_r^1$, $B'' \in \mathcal{T}_r^2$. Since r is a d-proximity, we have $cl_r^1 A' \, r \, cl_r^2 B'$. Now

$$cl_r^1 A = \bigcup \{cl_r^1 A' : A' \in \mathcal{A}\}, \quad cl_r^2 B = \bigcup \{cl_r^2 B' : B' \in \mathcal{B}\},$$

and r is a q-proximity, thus $cl_r^1 A \, r \, cl_r^2 B$. $\square$

7.2 REMARKS. a) The following could be added to 7.1:

(v) r is a wqd-proximity;

(vi) r is a qwd-proximity.

In (iii), "dq-proximity" can be replaced by "q-proximity".

b) Each quasi-proximity is a Leader proximity (4.15).

c) In 6.12, (ii) means that r is a d-proximity satisfying L* and r^q is a quasi-proximity. Here L* cannot be replaced by L. Indeed, let $X = \{1, 2, 3, 4, 5\}$ and let r be the coarsest symmetrical d-proximity on X for which $\{x\} \, r \, \{y\}$ $(x, y \in X, \ x \neq y)$ and $\{1, 2\} \, r \, \{3, 4\}$. r^q is a proximity (the discrete one), L* is fulfilled, but L does not hold ($c_r^1 \{1, 2\}$ and $c_r^2 \{3, 4\}$ are not even disjoint).

7.3 The w-proximities [symmetrical w-proximities] will serve as preproximities for Leader proximities [Lodato proximities]. To see the analogy with the case of the quasi-proximities, observe that the following conditions are equivalent for a disjointness relation r:

(i) r is a w-proximity;

(ii) r is an $(a_1 a_2, \, bdq, \, bdq)$-preproximity, where $a_1 = w$ or h, $a_2 = dq$ or h, $b = w$ or h;

(iii) $r \subset r^w$;

(iv) dom r is a subbase for $\mathcal{T}_r^1$, ran r is a subbase for $\mathcal{T}_r^2$.

(Compare (ii) with 5.7, (iii) with 5.27 a), (iv) with 5.26.) The results of § 6 remain valid with the following substitutions:

$$\mathsf{PrProx} \mapsto w\mathsf{Prox}$$
$$\mathsf{MeProx} \mapsto wd\mathsf{Prox}$$
$$\mathsf{QProx} \mapsto wdq\mathsf{Prox}$$
$$\mathsf{MeProx}^w \mapsto dw\mathsf{Prox}$$
$$\mathsf{QProx}^w \mapsto dqw\mathsf{Prox}$$
$$e \mapsto wdq.$$

Moreover, each category in the left-hand column is bireflective in the category it is to be replaced by.

REFERENCES

[1] Aarts, J. M. and de Groot, J., Complete regularity as a separation axiom, *Canad. J. Math.* **21** (1969), No 1, 96—105. *MR* **38** # 5160.

[2] Čech, E., *Topological spaces,* Revised by Z. Frolík and M. Katětov, Academia, Prague and Interscience, London, 1966. *MR* **35** # 2254.

[3] Császár, Á., Sur une classe de structures topologiques générales, *Revue Math. Pures Appl.* **2** (1957), 399—407. *MR* **20** # 1289.

[4] Császár, Á., *Fondements de la topologie générale,* Akadémiai Kiadó, Budapest, 1960. *MR* **22** # 4043.

[5] Császár, Á., *Foundations of general topology,* Pergamon Press, Oxford, 1963. *MR* **28** # 575.

[6] Császár, Á., On the characterization of completely regular spaces, *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.* **11** (1968), 79—82. *MR* **39** # 6247.

[7] Császár, Á., Doppeltkompakte bitopologische Räume, *Theory of sets and topology* (in honour of Felix Hausdorff), VEB Deutscher Verlag Wissensch., Berlin, 1972, 59—67. *MR* **49** # 7990

[8] Császár. Á., *General topology,* Akadémiai Kiadó, Budapest and Adam Hilger Ltd, Bristol, 1978. *MR* **57** # 13812.

[9] Császár, Á., Proximities, screens, merotopies, uniformities I, *Acta Math. Hungar.* **49** (1987), 459—479. *MR* **88k**: 54002a.

[10] Deák, E., Eine vollständige Charakterisierung der Teilräume eines euklidischen Raumes mittels der Richtungsdimension, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **9** (1964), No 3, 437—464. *MR* **32** # 4663b.

[11] Deák, E. and Hamburger, P., Vollständig interne Charakterisierungen der T_2-kompaktifizierbaren Räume, *Period. Math. Hungar.* **4** (1973), No 2—3, 125—145. *MR* **50** # 14654.

[12] Deák, J., On bitopological spaces I, *Studia Sci. Math. Hungar.* (to appear).

[13] Deák, J., On bitopological spaces III, *Studia Sci. Math. Hungar.* (to appear).

[14] Dvalishvili, B. P., Separation in bitopological spaces, *Sakharth. SSR Mecn. Akad. Moambe = Soobshch. Akad. Nauk Gruzin. SSR* **73** (1974), No 2, 285—288 (in Russian; English summary). *MR* **50** # 5757.

[15] Dvalishvili, B. P., The dimension and certain other questions of the theory of bitopological spaces, *Collection of papers on topology I, Dimension theory and sheaf theory, Sakharth. SSR Mecn. Akad. Math. Inst. Shrom = Trudy Tbiliss. Mat. Inst. Razmadze Akad. Nauk Gruzin. SSR* **56** (1977), 15—51 (in Russian; English summary). *MR* **58** # 18366.

[16] Efremovich, V. A., The geometry of proximity I, *Mat. Sbornik N. S.* **31** **(73)** (1952), 189—200 (in Russian). *MR* **14**—1106.

[17] Engelking, R., *General topology,* PWN, Warsaw, 1977. *MR* **58** # 18316b.

[18] Fletcher, P., Hoyle, H. B. III and Patty, C. W., The comparison of topologies, *Duke Math. J.* **36** (1969), No 2, 325—331. *MR* **39** # 3441.

[19] Fletcher, P. and Lindgren, W. F., Note on a result of E. P. Lane, *Portugal. Math.* **35** (1976), No 3—4, 211—212. *MR* **57** # 17597.

[20] Frink, O., Compactifications and seminormal spaces, *Amer. J. Math.* **86** (1964), No 3, 602—607. *MR* **29** # 4028.

[21] Gastl, G. C., Bitopological spaces from quasiproximities, *Portugal. Math.* **33** (1974), No 3—4, 213—218. *MR* **52** # 6668.

[22] Hamburger, P., Complete regularity as a separation axiom, *Period. Math. Hungar.* **4** (1973), No 1, 39—44. *MR* **49** # 6144.

[23] Hamburger, P., A general method to give internal characterizations of completely regular and Tychonoff spaces, *Acta Math. Acad. Sci. Hungar.* **23** (1972), No 3—4, 479—494. *MR* **47** # 5829.

[24] Hamburger, P., On internal characterizations of complete regularity and Wallman-type compactifications, *General topology and its relations to modern analysis and algebra III* (Proc. Third Prague Topological Sympos., 1971), Academia, Prague and Academic Press, New York, 1972, 171—172. *MR* **49** # 3800.

[25] Hamburger, P., Remarks on the lecture "Zamečanie po povodu ponjatia okrestnosti v topologii" given by P. S. Aleksandrov, *Topology and its applications* (Proc. Internat. Sympos. Topology Appl., Budva, 1972), Savez Društva Mat. Fiz. Astron., Belgrade, 1973, 106—112. *MR* **49** # 3835.

[26] Hamburger, P., Superextensions and preproximities, *Topics in topology* (Proc. Third Colloq.,

Keszthely, 1972), Colloq. Math. Soc. J. Bolyai **8**, North-Holland, Amsterdam, 1974, 367—378. *MR* **51** # 6730.

[27] HAMBURGER, P., Előszomszédsági relációk és alkalmazásuk a bővítéselméletben, Dissertation, Budapest, 1972.

[28] HERRLICH, H., Topological structures, *Topological structures* (Proc. Sympos. in honour of Johannes de Groot, Amsterdam, 1973), Math. Centre Tracts **52**, Mathematisch Centrum, Amsterdam, 1974, 59—122. *MR* **50** # 11165.

[29] HERRLICH, H., Categorical topology 1971—1981, *General topology and its relations to modern analysis and algebra V* (Proc. Fifth Prague Topological Sympos., 1981), Sigma Series in Pure Math. **3**, Heldermann, Berlin, 1983, 279—386. *MR* **84d**: 54016.

[30] HERRLICH, H. and STRECKER, G. E., *Category theory: an introduction*, Allyn and Bacon, Boston, 1973. *MR* **50** # 2284.

[31] HUNSAKER, W. and LINDGREN, W., Construction of quasi-uniformities, *Math. Ann.* **188** (1970), No 1, 39—42. *MR* **42** # 1057.

[32] JUHÁSZ, I. and MACHOVER, M., A note on non-standard topology, *Proc. Kon. Ned. Akad. Wetensch.* **72** = *Indag. Math.* **31** (1969), No 5, 482—484. *MR* **40** # 6488.

[33] LANE, E. P., Bitopological spaces and quasi-uniform spaces, *Proc. London Math. Soc.* (3) **17** (1967), No 2, 241—256. *MR* **34** # 5054.

[34] LANE, E. P., Quasi-proximities and bitopological spaces, *Portugal. Math.* **28** (1968), No 3—4, 151—159. *MR* **44** # 3284.

[35] LEADER, S., On products of proximity spaces, *Math. Ann.* **154** (1964), No 2, 185—194. *MR* **28** # 5420.

[36] LESEBERG, D., Neighbourhood structures, *Topology* (Proc. Fourth Colloq., Budapest, 1978), Vol. II, Colloq. Math. Soc. J. Bolyai **23**, North-Holland, Amsterdam, 1980, 751—804. *MR* **82b**: 54018.

[37] LODATO, M. W., On topologically induced general proximity relations, *Proc. Amer. Math. Soc.* **15** (1964), No 3, 417—422. *MR* **28** # 4513.

[38] MURDESHWAR, M. G. and NAIMPALLY, S. A., *Quasi-uniform topological spaces*, Nordhoff, Groningen, 1966. *MR* **35** # 2267.

[39] NAIMPALLY, S. A. and WARRACK, B. D., *Proximity spaces*, Cambridge Univ. Press, London, 1970. *MR* **43** # 3992.

[40] PERVIN, W. J., Quasi-proximities for topological spaces, *Math. Ann.* **150** (1963), 325—326. *MR* **27** # 2767.

[41] SAEGROVE, M. J., Pairwise complete regularity and compactification in bitopological spaces, *J. London Math. Soc.* (2) **7** (1973), No 2, 286—290. *MR* **49** # 11483.

[42] SALBANY, S., *Bitopological spaces, compactifications and completions*, Math. Monographs Univ. Cape Town **1**, Department Math., Univ. Cape Town, Cape Town, 1974. *MR* **54** # 13869.

[43] SHARMA, P. L., Proximity bases and subbases, *Pacific J. Math.* **37** (1971), No 2, 515—526. *MR* **46** # 4488.

[44] SMIRNOV, YU. M., On the theory of completely regular spaces, *Dokl. Akad. Nauk SSSR* **62** (1948), No 6, 749—752 (in Russian). *MR* **10**—315.

[45] STEEN, L. A. and SEEBACH, J. A., JR., *Counterexamples in topology*, Holt, Rinehart and Winston, New York, 1970. *MR* **42** # 1040.

[46] STEINER, E. F., The relation between quasi-proximities and topological spaces, *Math. Ann.* **155** (1964), 194—195. *MR* **29** # 581.

[47] STEINER, E. F., Normal families and completely regular spaces, *Duke Math. J.* **33** (1966), No 4, 743—745. *MR* **33** # 7975.

[48] WALLACE, A. D., Separation spaces, *Ann. of Math.* **42** (1941), No 3, 687—689. *MR* **3**—57.

[49] WATTEL, E., Subbase structures in nearness spaces, *General topology and its relations to modern analysis and algebra IV* (Proc. Fourth Prague Topological Sympos., 1976), Part B, Soc. Czechoslovak Math. Phys., Prague, 1977, 500—505. *MR* **57** # 7499.

[50] ZAĬTSEV, V. I., On the theory of Tikhonov spaces, *Vestnik Moskov. Univ. Ser I Mat. Mekh.* **22** (1967), No 3, 48—57 (in Russian; English summary). *MR* **35** # 7286.

[51] AARTS, J. M. and MRŠEVIĆ, M., Pairwise complete regularity as a separation axiom, *Reports of the Faculty of Mathematics and Information* 87—57, Delft, 1987.

(Received May 4, 1986)

MTA MATEMATIKAI KUTATÓ INTÉZETE
P.O. BOX 127
H—1364 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 179—186

VARIATIONS ON TIGHTNESS

I. JUHÁSZ

Abstract

Let $T(X)$ be the smallest cardinal τ such that whenever $\{F_\alpha: \alpha\in\varrho\}$ is an increasing sequence of closed sets in X with $\tau<cf(\varrho)$ then $\bigcup_{\alpha\in\varrho} F_\alpha$ is closed as well. $t_s(X)$ is the smallest cardinal $\varkappa$ such that whenever $p\in\bar{A}$ then there is a family $\{B_\alpha: \alpha\in\varkappa\}$ of subsets of A such that $p\in\overline{\bigcup_{\alpha\in\varkappa} B_\alpha}$ but $p\notin\bigcup_{\alpha\in\varkappa}\bar{B}_\alpha$. We show that $t_s(X)\leqq T(X)\leqq t(X)$ and give conditions for equality here as well as examples for inequality. We study the problem from [2] whether $t(X)=t_s(X)$ for a chain-net space X. We show that for such an X, $T(X)=t_s(X)$, give conditions for $t(X)=t_s(X)$ and then present two counterexamples.

The atomic tightness $a(p, A)$ of a limit point p of a set A in a space X is defined as follows:

$$a(p, A) = \min\{|B|: B\subset A\setminus\{p\} \text{ and } p\in\bar{B}\}.$$

The tightness $t(p, X)$ of a (non-isolated) point p of X is then obtained as

$$t(p, X) = \sup\{a(p, A): p\in A'\},$$

and the tightness of X is $t(X)=\sup\{t(p, X): p\in X'\}$.

As a trivial consequence of this definition we obtain the following property of tightness: If $\{F_\alpha: \alpha\in\varrho\}$ is an increasing sequence of closed sets in X and $t(X)<\mathrm{cf}(\varrho)$ then $\bigcup_{\alpha\in\varrho} F_\alpha$ is also closed. This leads us to the following definition.

Definition 1. For any space X let $T(X)$ denote the smallest cardinal τ with the property that whenever $\{F_\alpha: \alpha\in\varrho\}$ is an increasing sequence of closed sets in X with $\tau<cf(\varrho)$ then $\bigcup_{\alpha\in\varrho} F_\alpha$ is closed as well.

This is the first variation on $t(X)$; clearly $T(X)\leqq t(X)$. Since many applications of $t(X)$ actually use $T(X)$, it is natural to raise the question whether $T(X)$ is the same as $t(X)$? While the answer to this question is "no", it will also turn out that in many important cases it is "yes".

Lemma 1. *Whenever $A\subset X$ and $p\in A'$ we have*

$$\mathrm{cf}(a(p, A)) \leqq T(X).$$

Research supported by Hungarian National Science Foundation Grant No. 1805.

1980 *Mathematics Subject Classification* (1985 Revision). Primary 54A25, 54A20; Secondary 03E35.

Key words and phrases. Tightness, chain-net space.

PROOF. Let $\varrho=a(p, A)$ and $B=\{q_\alpha: \alpha\in\varrho\}$ be such that $B\subset A\setminus\{p\}$ and $p\in\overline{B}$. Then $\{F_\alpha=\overline{\{q_\beta: \beta\in\alpha\}}: \alpha\in\varrho\}$ is an increasing sequence of closed sets with $F=\bigcup_{\alpha\in\varrho} F_\alpha$ non-closed, since $p\notin F\supset B$, hence indeed $\mathrm{cf}(\varrho)\leqq T(X)$.

As an immediate corollary we obtain the following result.

COROLLARY 1. *If $t(X)=\lambda^+$ is a successor cardinal then $t(X)=T(X)$.*

The next corollary will be used later.

COROLLARY 2. *If in X every $<\aleph_\omega$-closed set is also closed then $t(X)=T(X)$. ($F\subset X$ is called $<\aleph_\omega$-closed if for every $A\subset F$ with $|A|<\aleph_\omega$ we have $\overline{A}\subset F$.)*

PROOF. Note first that our assumption clearly implies $t(X)\leqq\aleph_\omega$, hence, by Corollary 1, we may as well assume that $t(X)=\aleph_\omega$.

Now, suppose that $T(X)=\aleph_n<t(X)=\aleph_\omega$. Then, by Lemma 1, we must have $a(p, A)\leqq\aleph_n$ or $a(p, A)=\aleph_\omega$ whenever $p\in A'$ in X, moreover p and A with $a(p, A)=\aleph_\omega$ must exist. Let us put

$$F=\bigcup\{\overline{B}: B\subset A \text{ and } |B|\leqq\aleph_n\},$$

we claim that F is $<\aleph_\omega$-closed. Indeed, if $H\subset F$ and $|H|<\aleph_\omega$ then by our above remark for every $q\in H'$ we must have $a(q, H)\leqq\aleph_n$ which immediately implies that $q\in F$. This, however, leads us to a contradiction since $p\in\overline{F}\setminus F$ shows that F is not closed.

PROPOSITION 1. *If X is compact T_2 then $t(X)=T(X)$.*

PROOF. Assume $T(X)<t(X)$, then as is well-known, see e.g. [1] or [4], X contains a free sequence $S=\{p : \alpha\in\varrho\}$ of length $\varrho=T(X)^+$. But then $\{F_\alpha=\overline{\{p_\beta: \beta\in\alpha\}}: \alpha\in\varrho\}$ is an increasing sequence of closed sets with $F=\bigcup_{\alpha\in\varrho} F_\alpha$ non-closed (indeed, no complete accumulation point of S belongs to F), which is a contradiction.

Now we want to show that $T(X)<t(X)$ is possible, even for "good" spaces X. Let $\varkappa$ be an arbitrary singular cardinal, i.e. $\sigma=\mathrm{cf}(\varkappa)<\varkappa$, we shall first produce a space $X(\varkappa)$ with a single nonisolated point such that $t(X(\varkappa))=\varkappa$ and $T(X(\varkappa))=\sigma$. The underlying set of $X(\varkappa)$ is $\varkappa+1$ and all points $\xi\in\varkappa$ are isolated in $X(\varkappa)$. To define the neighbourhoods of $\varkappa$ we first fix an increasing sequence $\{\varkappa_\gamma: \gamma\in\sigma\}$ of regular cardinals converging up to $\varkappa$. Then a set $A\cup\{\varkappa\}\subset\varkappa+1$ is a neighbourhood of $\varkappa$ in $X(\varkappa)$ iff there is some $\mu\in\sigma$ with the property that

$$|\varkappa_\nu\setminus A|<\varkappa_\nu$$

whenever $\nu\in\sigma\setminus\mu$.

Now, we show that $t(X(\varkappa))=t(\varkappa, X(\varkappa))=\varkappa$. Indeed, if $A\subset\varkappa$ and $|A|<\varkappa$ then we must have a $\mu\in\sigma$ with $|A|<\varkappa_\mu$, hence $|A\cap\varkappa_\nu|<\varkappa_\nu$ for all $\nu\in\sigma\setminus\mu$. Clearly, then $\varkappa\notin\overline{A}$.

Next, to show $T(X)\leqq\sigma$, consider any strictly increasing sequence $\{F_\alpha: \alpha\in\varrho\}$ of sets closed in $X(\varkappa)$ with $\varrho=cf(\varrho)>\sigma$. If $\varkappa\in F_\alpha$ for some α then of course

$F=\bigcup_{\alpha\in\varrho} F_\alpha$ is trivially closed, hence we may assume that $\varkappa\notin F$. Note that since $\varrho\leq\varkappa$ and ϱ is regular we actually have $\varrho<\varkappa$, hence $\varrho<\varkappa_{\bar\mu}$ for some $\bar\mu\in\sigma$.

Since for each $\alpha\in\varrho$ the set F_α is closed and $\varkappa\notin F_\alpha$ we may define $\mu_\alpha\in\sigma$ by

$$\mu_\alpha = \min\{\mu_\beta\colon\ |F_\alpha\cap\varkappa_\nu| < \varkappa_\nu \ \text{ for all } \ \nu\in\sigma\setminus\beta\}.$$

Since the F_α's are increasing we have that $\alpha<\alpha'$ implies $\mu_\alpha\leq\mu_{\alpha'}$, consequently as ϱ is regular with $\sigma<\varrho<\varkappa$ we get that the sequence $\{\mu_\alpha\colon\ \alpha\in\varrho\}$ is bounded in σ. Thus we may choose an ordinal $\mu\in\sigma$ such that $\bar\mu\leq\mu$ and $\mu_\alpha\leq\mu$ for all $\alpha\in\varrho$. But then for each $\nu\in\sigma\setminus\mu$ the regularity of $\varkappa_\nu$ implies that

$$|F\cap\varkappa_\nu| = \Big|\bigcup_{\alpha\in p} F_\alpha\cap\varkappa_\nu\Big|\leq \sum_{\alpha\in\varrho}|F_\alpha\cap\varkappa_\nu| < \varkappa_\nu,$$

showing that $\varkappa\notin\bar F$, hence F is closed.

We may now give a result that may be contrasted with Corollary 1.

PROPOSITION 2. *For every limit cardinal λ there is a space X with $t(X)=\lambda$ and $T(X)=\aleph_0$.*

PROOF. If $\mathrm{cf}(\lambda)=\aleph_0$ then of course we may take $X(\lambda)$ as our X. Otherwise we may note that $\lambda=\bigcup\{\varkappa<\lambda\colon\ \mathrm{cf}(\varkappa)=\aleph_0\}$, consequently if we put

$$X = \oplus\{X(\varkappa)\colon\ \varkappa<\lambda \ \text{and}\ \mathrm{cf}(\varkappa)=\aleph_0\},$$

i.e. X is the topological sum of these $X(\varkappa)$, then we again have $t(X)=\lambda$ and $T(X)=\aleph_0$, which is checked easily. Note that the $X(\varkappa)$ are "very good" spaces, having only one non-isolated point, e.g. they are hereditarily paracompact, hence so is X.

Next we turn to the second variation of $t(X)$, that we call $t_s(X)$, the set tightness of X. This notion was originally introduced in [2], where, rather unjustifiably, it was called quasi character.

DEFINITION 2. If $p\in A'$ then the atomic set tightness $a_s(p, A)$ is defined as follows:

$$a_s(p, A) = \min\big\{|\mathscr{B}|\colon\ \mathscr{B}\subset P(A\setminus\{p\}) \ \ \&\ \ p\in\overline{\cup\mathscr{B}} \ \ \&\ \ p\notin\cup\{\bar B\colon\ B\in\mathscr{B}\}\big\}.$$

Then $t_s(p, X)=\sup\{a_s(p, A)\colon\ p\in A'\}$ and $t_s(X)=\sup\{t_s(p, X)\colon\ p\in X\}=$ $=\sup\{a_s(p, A)\colon\ p\in X\ \&\ p\in A'\}$.

It is obvious that $t_s(X)\leq t(X)$, and in fact we shall show below that for any X we have $t_s(X)\leq T(X)$. Before we do this, however, we formulate a lemma that lists some interesting properties of the atomic set tightness $a_s(p, A)$.

LEMMA 2. (i) *$a_s(p, A)$ is equal to the minimum number of neighbourhoods of p in $A\cup\{p\}$ whose intersection is not a neighbourhood of p in $A\cup\{p\}$; in particular, in a T_1 space $a_s(p, A)\leq\psi(p, A\cup\{p\})$.*

(ii) *$a_s(p, A)$ is always a regular cardinal.*

(iii) *If X is regular then $a_s(p, A)=a_s(p, \bar A)$.*

PROOF. (i) is clearly just a reformulation of the definition of $a_s(p, A)$.

(ii) Let $a_s(p, A)=\varkappa$ and assume that $\mathcal{B}\subset P(A)$ witnesses this. If $\varkappa$ were singular then we could write

$$\mathcal{B} = \bigcup\{\mathcal{B}_\nu\colon\ \nu\in \mathrm{cf}(\varkappa)\}$$

with $|\mathcal{B}_\nu|<\varkappa$ for all ν. But then, by definition, $p\notin \overline{\bigcup\mathcal{B}_\nu}$ for $\nu\in\mathrm{cf}(\varkappa)$, hence the family $\{\bigcup\mathcal{B}_\nu\colon\ \nu\in\mathrm{cf}(\varkappa)\}$ would show that $a_s(p, A)\leqq\mathrm{cf}(\varkappa)<\varkappa$, a contradiction.

(iii) Now assume that $a_s(p, \bar{A})=\varkappa$ and let $\mathcal{B}\subset P(\bar{A})$ be a witness for this, with $B\neq\emptyset$ for all $B\in\mathcal{B}$. Since $p\notin\bar{B}$ for $B\in\mathcal{B}$ we have, by the regularity of X disjoint open sets $U_B\ni p$ and $V_B\supset\bar{B}$. We claim that the family $\mathcal{C}=\{A\cap V_B\colon\ B\in\mathcal{B}\}$ shows $a_s(p, A)\leqq\varkappa$. Indeed, if G is any open set containing p we have some $B\in\mathcal{B}$ with $G\cap B\neq\emptyset$ which, by $B\subset\bar{A}$ and $B\neq\emptyset$, clearly implies $G\cap V_B\cap A\neq\emptyset$ as well, hence $p\in\overline{\bigcup\mathcal{C}}$. On the other hand for every $V_B\cap A\in\mathcal{C}$ we have that $p\notin\overline{V_B\cap A}$ since $U_B\cap V_B=\emptyset$. Since $a_s(p, \bar{A})\leqq a_s(p, A)$ holds trivially we indeed have $a_s(p, A)=a_s(p, \bar{A})$.

COROLLARY 3. *If X is T_2 then $t_s(X)\leqq s(X)$, see* [2], *if X is T_3 then $t_s(X)\leqq F(X)$, see* [3].

PROOF. Let $s(X)=\varkappa$, then if X is T_2 for any p and A in X we have (see [4] or [8]) that either $a(p, A)\leqq\varkappa$ or $\psi(p, A\cup\{p\})\leqq\varkappa$, hence, by (i), we must have $a_s(p, A)\leqq\varkappa$. To see the second half now assume that X is T_3. Then $a_s(p, A)=a_s(p, \bar{A})$ by (iii), moreover using (i) we may define a strongly decreasing sequence $\{V_\alpha\colon\ \alpha\in\varkappa\}$ of open neighbourhoods of p in $\bar{A}$, i.e. such that $\overline{V_\alpha}\subset V_\beta$ if $\beta<\alpha$. Now if $p_\alpha\in V_\alpha\setminus\overline{V_{\alpha+1}}$ for all $\alpha\in\varkappa$ then $\{p_\alpha\colon\ \alpha\in\varkappa\}$ is a free sequence in $\bar{A}$ and thus in X as well.

COROLLARY 4. *For any X we have $t_s(X)\leqq T(X)$.*

PROOF. Let $a_s(p, A)=\varkappa$ and let $\mathcal{B}\subset P(A)$ witness this. Note that, by (ii), $\varkappa$ is regular. Write $\mathcal{B}=\{B_\alpha\colon\ \alpha\in\varkappa\}$ and for each $\alpha\in\varkappa$ let $F_\alpha=\overline{\bigcup\{B_\beta\colon\ \beta\in\alpha\}}$. Then $\{F_\alpha\colon\ \alpha\in\varkappa\}$ is an increasing sequence of closed sets of length $\varkappa=\mathrm{cf}(\varkappa)$ whose union is not closed because it does not contain p. Consequently, we have $a_s(p, A)\leqq T(X)$ for all p and A in X, hence $t_s(X)\leqq T(X)$.

Now we give examples which show that in general $t_s(X)$ can be less than $T(X)$.

PROPOSITION 3. *For every cardinal $\varkappa$ there is a (very good) space $X_\varkappa$ with $T(X_\varkappa)=\varkappa$ and $t_s(X_\varkappa)=\aleph_0$.*

PROOF. Let us first assume that $\varkappa$ is regular and define $X_\varkappa$ on the underlying set $(\varkappa\times\omega)\cup\{p\}$ as follows: Every point in $\varkappa\times\omega$ is isolated and a neighbourhood for p consists of the sets of the form

$$(\varkappa\setminus\alpha)\times(\omega\setminus n)\cup\{p\}$$

where $\alpha\in\varkappa$ and $n\in\omega$. It is straightforward to check that $T(X_\varkappa)=\varkappa$ and $t_s(X_\varkappa)=\aleph_0$. Next if λ is singular, then we may put $X_\lambda=\oplus\{X_\varkappa\colon\ \varkappa<\lambda\ \&\ \varkappa=\mathrm{cf}(\varkappa)\}$.

The rest of this paper will be devoted to the following problem raised in [2]: Does $t(X)=t_s(X)$ hold in a chain-net space X? (Recall that a space is chain-net if every non-closed set in it contains a well ordered sequence — or chain-net — that converges out of it.)

PROPOSITION 4. *In a chain-net space X we have $T(X)=t_s(X)$.*

PROOF. Let $\{F_\alpha: \alpha\in\varrho\}$ be an increasing sequence of closed sets in X such that $F=\bigcup_{\alpha\in\varrho} F_\alpha$ is non-closed and $\varrho=\mathrm{cf}(\varrho)$. Since X is chain-net, we may then find a well-ordered sequence $\{x_\nu: \nu\in\mu\}\subset F$ converging to a point $x\in X\setminus F$. Again, we may assume that $\mu=\mathrm{cf}(\mu)$. We claim that then $\mu=\varrho$. Indeed, we cannot have $\mu<\varrho$ since then, by the regularity of ϱ, we had $\{x_\nu: \nu\in\mu\}\subset F_\alpha$ for some α, hence $x\in F_\alpha\subset F$, a contradiction. On the other hand $\varrho<\mu$ would imply that $|\{\nu\in\mu: x_\nu\in F_\alpha\}|=\mu$ holds for some $\alpha\in\varrho$, which again would imply $x\in F_\alpha\subset F$, a contradiction.

Let us put $A=\{x_\nu: \nu\in\varrho\}$, we claim that $a_s(x, A)=\varrho$. In fact this follows from the fact that if $B\subset A$ and $|B|<\varrho$ then $B\subset F_\alpha$ for some α. This shows that we have $T(X)\leqq t_s(X)$ because if $\{F_\alpha: \alpha\in\varrho\}$ is an increasing sequence of closed sets and $\mathrm{cf}(\varrho)>t_s(X)$ then $\bigcup_{\alpha\in\varrho} F_\alpha$ must be closed. But then, by Proposition 3, we have $T(X)=$ $=t_s(X)$.

COROLLARY 5. *Let X be a chain-net space. Then $t(X)=t_s(X)$ provided one of the following conditions holds:*

a) *$t(X)$ is a successor cardinal,*

b) *X is compact T_2,*

c) *X does not contain any convergent chain net whose length is a regular cardinal $\varrho>\aleph_\omega$.*

PROOF. a) and b) are immediate from Proposition 4 and Corollary 1 and Proposition 1, respectively. c) will follow from Corollary 2 if we can show that in this case every $<\aleph_\omega$-closed set F in X is closed. Indeed, otherwise some sequence of a regular length ϱ would converge out of F, where $\varrho<\aleph_\omega$ is impossible since F is $<\aleph_\omega$-closed and $\varrho>\aleph_\omega$ is impossible because of our assumption.

Note that condition c) is trivially satisfied if $\psi(X)\leqq\aleph_\omega$ or if $|X|\leqq\aleph_\omega$. (Of course, for T_1 spaces the second condition implies the first.) Consequently, a counter-example, i.e. a chain-net space X with $t(X)\neq t_s(X)$ must have cardinality $>\aleph_\omega$. This leads us to the following natural question:

PROBLEM 1. Does there exist a (T_2 or T_3) chain-net space X with ($|X|=\aleph_{\omega+1}$ and $t(X)\neq t_s(X)$?

In what follows and in [7] partial answers to this problem will be given.

PROPOSITION 5. *There exists a T_2 chain-net space X with $|X|=(c^{+\omega})^{\aleph_0}$ and $t(X)\neq t_s(X)$. (Here $c^{+\omega}$ denotes the ω^{th} successor of the continuum. Note that, e.g. under GCH, $(c^{+\omega})^{\aleph_0}$ may be equal to $\aleph_{\omega+1}$.)*

Proof. Let us write $c^{+\omega}=\varkappa$ and $(c^{+\omega})^{\aleph_0}=\lambda$. Our space X to be constructed will be of the form $Y\cup\{p\}$, so we start with constructing the space Y. Now Y will be a disjoint union $Y=\bigcup_{n\in\omega} Y_n$ where the Y_n's are defined by induction as follows: $Y_0=\lambda$ and if Y_n is defined then Y_{n+1} is chosen to be a maximal almost disjoint subfamily of $[Y_n]^{\aleph_0}$. We shall say that $q\in Y=\bigcup_{n\in\omega} Y_n$ has rank n (in symbols: $\operatorname{rk}(q)=n$) if $q\in Y_n$.

A topology is defined on Y by declaring $G\subset Y$ to be open iff for every $q\in G$ with $\operatorname{rk}(q)>0$ there is a finite subset $F\subset q$ such that $q\setminus F\subset G$.

Next to show that Y is T_2 in this topology we first prove the following claim: if A is any countable subset of Y_n then the points in A can be separated by disjoint open sets. For $n=0$ this is trivial, so assume that we know it up to n and let $A\subset Y_{n+1}$ be countable. Since the elements of A are almost disjoint, it is easy to fix for each $q\in A$ a finite subset $F_q\subset q$ such that $\{q\setminus F_q\colon q\in A\}$ is disjoint. Applying the claim to $\bigcup\{q\setminus F_q\colon q\in A\}\subset Y_n$ we may now easily get the required separation of A by open sets.

Now to show that Y is T_2, consider two points $q_1, q_2\in Y$. If $\operatorname{rk}(q_1)=\operatorname{rk}(q_2)$ then q_1 and q_2 can be separated by the claim. Thus assume that $\operatorname{rk}(q_1)=n_1<\operatorname{rk}(q_2)=n_2$. Next we may choose a countable neighbourhood V_2 of q_2 such that $q_1\notin V_2$, and apply the claim to the countable set $\{q_1\}\cup(V_2\cap Y_{n_1})\subset Y_{n_1}$ to obtain disjoint open sets around q_1 and the points in $V_2\cap Y_{n_1}$. It is obvious that in this way we can obtain disjoint neighbourhoods of q_1 and q_2.

Finally we show that Y is sequential, hence chain-net. Thus let $A\subset Y$ be non-closed and let q have minimal rank in $\bar{A}\setminus A$. We claim that $q\cap A$ is infinite and hence yields a sequence converging to $q\notin A$. Indeed, for every $s\in q\setminus A$ $\operatorname{rk}(s)<\operatorname{rk}(q)$ implies $s\notin\bar{A}$ and thus an open set G_s containing s may be chosen with $A\cap G_s=\emptyset$, hence if $q\cap A$ were finite then

$$\bigcup\{G_s\colon s\in q\setminus A\}\cup\{q\}$$

would be a neighbourhood of q that misses A, a contradiction.

Now we add a new point p to Y to obtain $X=Y\cup\{p\}$ and we need to define the neighbourhoods of p in X. To this end we shall say that a subset A of Y is small if $|A\cap Y_n|<\varkappa$ holds for all $n\in\omega$, and we denote by $\mathscr{F}$ the family of all small closed sets in Y. Then a neighbourhood base for p in X is given by the sets $X\setminus F$ with $F\in\mathscr{F}$. It is straightforward to see that this way we get a T_2 topology on X.

In order to establish the required properties of X we need the following facts.

Fact 1. If $A\subset Y$ is small so is $\bar{A}$. Since, as is easily seen, $\bar{A}=\bigcup_{n\in\omega}\overline{A\cap Y_n}$ and consequently $\bar{A}\cap Y_k=\bigcup\{\overline{A\cap Y_n}\colon n\le k\}\cap Y_k$, it clearly suffices to show that if $A\subset Y_k$ and $|A|<\varkappa$ then $|\bar{A}|<\varkappa$ as well. However, in a sequential space we have for any set A that $|\bar{A}|\le|A|^{\aleph_0}$, hence our claim follows from the fact that $(c^{+n})^{\aleph_0}=c^{+n}<\varkappa$ for all $n\in\omega$.

Fact 2. If $A\subset Y$ is not small then $|\bar{A}\cap Y_k|>\varkappa$ for some $k\in\omega$. Indeed, assume that $|A\cap Y_n|\ge\varkappa$; then the infinite members of $Y_{n+1}\restriction A\cap Y_n$ form a maximal almost disjoint family in $[A\cap Y_n]^{\aleph_0}$, which is well-known to be of size $>\varkappa$ (in fact of size equal to λ). On the other hand, clearly every member of Y_{n+1} that has infinite intersection with $A\cap Y_{n+1}$ is in the closure of A, i.e. $|\bar{A}\cap Y_{n+1}|>\varkappa$.

Now to see that X is chain-net consider any non-closed set $F\subset X$. We may of course assume that $F\cap Y$ is closed in Y since Y is sequential and, since F is non-closed in X we must have $F\notin\mathscr{F}$. But then, by Fact 2, we have $|F\cap Y_k|>\varkappa$ for some $k\in\omega$. It is clear, however, that any sequence $\{x_\nu\colon \nu\in\varkappa^+\}\subset F\subset Y_k$ consisting of pairwise distinct points converges (out of F) to p.

Next, by Fact 1, for any $A\subset Y$ with $|A|<\varkappa$ we have $p\notin\bar{A}$, hence $t(p,X)\geq a(p,X)\geq\varkappa$. Finally, if $A\subset Y$ is such that $p\in\bar{A}$ then again by Fact 1 we have $|A\cap Y_k|\geq\varkappa$ for some $k\in\omega$. Thus we may choose for each $n\in\omega$ a set $B_n\subset A\cap Y_k$ with $|B_n|=c^{+n}$. Clearly, we have $p\notin\bar{B}_n$ for all $n\in\omega$, but $p\in\overline{\bigcup_{n\in\omega}B_n}$ since $\left|\bigcup_{n\in\omega}B_n\right|=\varkappa$. This shows that $a_s(p,A)=\aleph_0$ for all such A, hence $t_s(p,X)=\aleph_0$ as well. But $t(q,X)=t(q,Y)=\aleph_0$ for every $q\in Y$ because Y is sequential, hence indeed $t_s(X)=\aleph_0$.

There are two things we do not like about the above example. First, its cardinality, $(c^{+\omega})^{\aleph_0}$, is only consistently equal to $\aleph_{\omega+1}$. Second, it is not regular. A regular example is constructed (in ZFC!) in [7], which has cardinality potentially even larger than $(c^{+\omega})^{\aleph_0}$, namely $\mathfrak{B}_{\omega+1}$. This, however, could still be equal to $\aleph_{\omega+1}$. Below we give a different T_3 example, which, however, is not obtained, and probably not obtainable, in ZFC.

Proposition 6. *Assume that there exists a countably compact, locally countable T_3 space Y with $|Y|=\aleph_{\omega+1}$ such that either* 1) $c<\aleph_\omega$ or 2) *Y is $\aleph_0$-fair. (Y is $\varkappa$-fair if whenever $A\subset Y$ and $|A|\leq\varkappa$ then $|\bar{A}|\leq\varkappa$ as well. It is known, cf.* [5] *or* [6], *that our assumption is satisfied under numerous set-theoretic hypotheses.) Then there is a 0-dimensional T_2 (hence T_3) chain-net space X with $|X|=\aleph_{\omega+1}$ and $t(X)\neq t_s(X)$.*

Proof. Let us note first that Y is 0-dimensional. Moreover, if Y is $\aleph_0$-fair it is also $\aleph_n$-fair for all $n\in\omega$, and in the first case it is obviously $\aleph_n$-fair whenever $c\leq\aleph_n<\aleph_\omega$. Combining this with local countability we get that every subset $A\subset Y$ with $|A|<\aleph_\omega$ can be covered with a clopen set Z with $|Z|<\aleph_\omega$.

Let us now put $X=Y\cup\{p\}$ where a neighbourhood base of p consists of sets of the form $X\setminus Z$ with $Z\subset Y$ clopen and $|Z|<\aleph_\omega$. Clearly, X is a 0-dimensional T_2 space.

We claim that X is also chain-net. To see this it suffices to consider a set $F\subset Y$ closed in Y and such that $p\in\bar{F}$, i.e. $|F|\geq\aleph_\omega$. Now, by [5], we actually must have $|F|>\aleph_\omega$, hence a one-to-one sequence $\{x_\nu\colon \nu\in\aleph_{\omega+1}\}\subset F$ can be chosen, which obviously converges to p.

Now, if $A\subset Y$ and $|A|<\aleph_\omega$ then a clopen set $Z\supset A$ exists with $|Z|<\aleph_\omega$, hence $p\notin\bar{A}$, showing that $t(X)\geq t(p,X)\geq\aleph_\omega$. Finally, if $A\subset Y$ and $p\in\bar{A}$ then we must have $|A|\geq\aleph_\omega$, hence we may choose sets $B_n\subset A$ with $|B_n|=\aleph_n$ for all $n\in\omega$. The family $\{B_n\colon n\in\omega\}$ clearly shows that $a_s(p,A)=\aleph_0$, hence we have $t_s(X)=\aleph_0$, as claimed.

REFERENCES

[1] Arhangel'skiĭ, A., Bicompacta that satisfy the Suslin condition hereditarily. Tightness and free sequences, *Dokl. Akad. Nauk SSSR* **199** (1971), 1227—1230. *MR* **44** # 5914.

[2] Archangelskiĭ, A., Isler, R. and Tironi, G., On pseudoradial spaces, *Comment. Math. Univ. Carolin.* **27**(1986), 137—154.

[3] Bella, A., Free sequences in pseudo-radial spaces, *Comment. Math. Univ-Carolin.* **27** (1986), 163—170.

[4] Juhász, I., *Cardinal functions in topology — ten years later*, Second edition, Mathematical Centre Tracts, 123, Mathematisch Centrum, Amsterdam, 1980. *MR* **82a**: 54002.

[5] Juhász, I., Nagy, Zs. and Weiss, W., On countably compact, locally countable spaces, *Period. Math. Hungar.* **10** (1979), 193—206. *MR* **80h**: 54025.

[6] Juhász, I., Shelah, S. and Soukup, L., More on countably compact, locally countable spaces, *Israel J. Math.* **62** (1988), 302—310.

[7] Juhász, I. and Weiss, W., On the tightness of chain-net spaces, *Comment. Math. Univ. Carolin.* **27** (1986), 677—681.

[8] Šapirovskiĭ, B. E., On discrete subspaces, *Soviet Math. Dokl.* **13** (1972), 215—219.

(Received June 5, 1986)

MTA MATEMATIKAI KUTATÓ INTÉZETE
P.O. BOX 127
H—1364 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 187—190

ON SOME DISTANCE PROPERTIES OF SETS OF POINTS IN GENERAL POSITION IN SPACE

ILONA PALÁSTI

Abstract

Is it possible to give n points in the space so that no three points are on a line, no four on a plane, no five on a sphere, if we assume moreover that they determine $n-1$ distinct distances among the possible $\binom{n}{2}$ ones such that the i-th distance occurs i times? Here we show that this is possible for $n=4, 5$ and 6.

P. Erdős [1] asked for n points in the plane, no three on a line, no four on a circle, such that they determine $n-1$ distinct distances, one occurring once, one twice, and so on, one $n-1$ times. The distances are not ordered by size or in any other way. Examples were given in [1] and [3] for $n=5$ and in [4] for $n=7$ and 6. Another example for $n=7$ was given in [5] and for $n=8$ in [6.] In this paper we consider the analogous problem in the space. Here we additionally suppose that no five points lie on a sphere. We have the following

THEOREM. *There exist six points in space no three on a line, no four in a plane, no five on a sphere, which determine 5 distinct distances d_i, $i=1, \dots, 5$, such that the distance d_i occurs i times.*

Let A, B, C and D be given so that

$$AB = AC = BC = 1 \quad \text{and} \quad AD = BD = CD = \sqrt{\frac{7}{12}}.$$

We choose the point E so that it lies on the same side of the plane ABC as D, and furthermore the segment BE is perpendicular to the plane ABC and has unit length (Fig. 1). Then we have

$$AB = AC = BC = BE = 1 \quad \text{and} \quad AE = CE = \sqrt{2}.$$

Let O denote the centre of the circumcircle of ABC. Then the line segment OD is parallel to BE and its length is $\frac{1}{2}$. This implies that the triangle BDE is isosceles.

Research (partly) supported by Hungarian National Foundation for Scientific Research Grant 1808.

1980 *Mathematics Subject Classification* (1985 Revision). Primary 51L15; Secondary 51M15, 51A20.

Key words and phrases. Arrangements of points, Erdős' problem.

Accordingly, the distance

$$AD = BD = CD = DE = \sqrt{\frac{7}{12}}$$

occurs four times.

Let F be a point such that it is separated from D and E by the plane ABC and the perpendicular from F to the plane ABC intersects this plane in a point, say S, which lies on the line segment AO. We shall show that it is possible to choose F so that

$$AF = 1 \quad \text{and} \quad DF = \sqrt{2}.$$

Writing $x = FS$ and $y = OS$ we have

$$AF^2 = x^2 + AS^2 = x^2 + \left(\frac{\sqrt{3}}{3} - y\right)^2$$

and

$$DF^2 = (FS + OD)^2 + y^2 = \left(x + \frac{1}{2}\right)^2 + y^2.$$

Thus the conditions $AF = 1$ and $DF = \sqrt{2}$ yield for x and y the equations

$$x^2 + y^2 - \frac{2\sqrt{3}}{3}y - \frac{2}{3} = 0$$

and

$$x^2 + y^2 + x - \frac{7}{4} = 0,$$

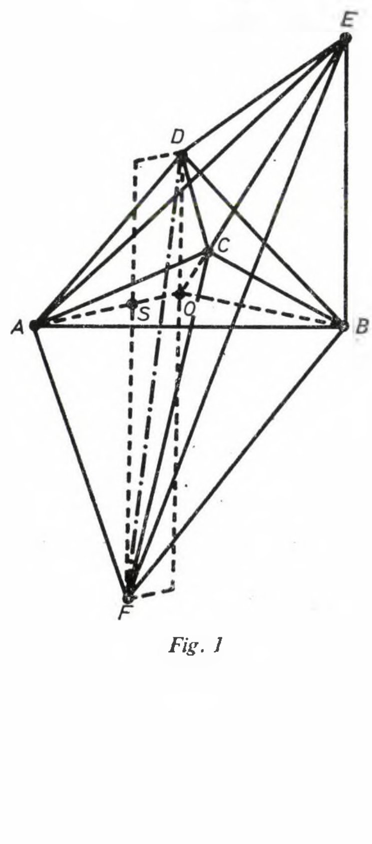

Fig. 1

or, which is the same

$$x = \frac{13}{12} - 2\frac{\sqrt{3}}{3}y, \quad \frac{7}{3}y^2 - \sqrt{3}\frac{19}{9}y + \frac{73}{144} = 0.$$

The last equations yield two positive solutions for y, by choosing the smaller one we have $y=0.1537...$ and $x=0.9058...$.

Now we are able to calculate the length of $BF=CF$ forming the two other edges of the tetrahedron $ABCF$. One easily obtains the equality

$$BS^2 = AB^2 + AS^2 - 2AB \cdot AS \cos 30^\circ = 1 + \left(\frac{\sqrt{3}}{3} - y\right)^2 - \sqrt{3}\left(\frac{\sqrt{3}}{3} - y\right) = 0.4457\ldots.$$

Since $FS=x$, therefore

$$BF^2 = x^2 + BS^2 = 1.2662\ldots.$$

Hence we obtain $BF=BC=1.1257...$.

It is easy to see that no three points from A, B, C, D, E and F are on a line, no four are on a plane, no five are on a sphere and the distance

$AB=AC=BC=BE=AF=1$	occurs five times,
$AD=BD=CD=DE=\sqrt{\frac{7}{12}}$	occurs four times,
$AE=CE=DF=\sqrt{2}$	occurs three times,
$BF=CF=1.1257...$	occurs twice,
$EF>BE+SF>1.9$	occurs once.

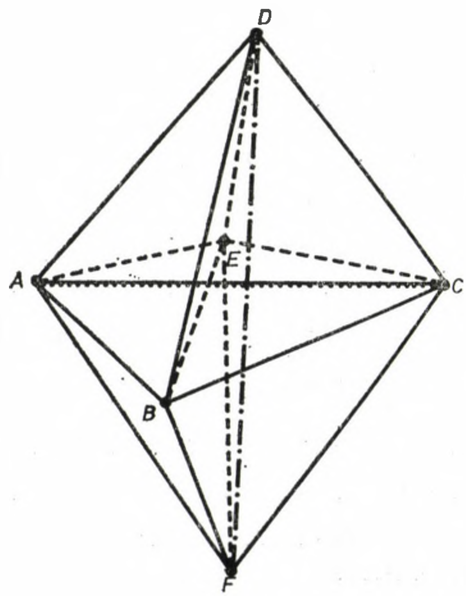

Fig. 2

The theorem is thus proved.

If we drop the point E then we have five points in the space such that the distances 1, $\sqrt{\frac{7}{12}}$, 1.1257... and $\sqrt{2}$ occur between these points four times, three times, twice and once, respectively.

Another set of six points with the property described in the Theorem can be constructed as follows: Let ABC and ACD be congruent equilateral triangles with the common side AC such that the planes ABC and ACD are perpendicular and

$$AB = BC = AC = AD = DC = 1.$$

Thus the distance 1 occurs five times among these four points and the sixth distance BD is equal to $\sqrt{\frac{3}{2}}$. We add to this set of points the centre E of the circumsphere of the tetrahedron $ABCD$. This gives rise to four more distances equal to $\frac{1}{2}\sqrt{\frac{5}{3}}$. One can show that a point F can be chosen in the normal bisector plane of the segment CD so that

$$FB = FC = BD = \sqrt{\frac{3}{2}} \quad \text{and} \quad FA = FE.$$

With appropriate coordinatization the points can be given as follows:

$$A = (0,0,0), \quad B = \left(\frac{\sqrt{3}}{2}, \frac{1}{2}, 0\right), \quad C = (0,1,0), \quad D = \left(0, \frac{1}{2}, \frac{\sqrt{3}}{2}\right),$$

$$E = \left(\frac{\sqrt{3}}{6}, \frac{1}{2}, \frac{\sqrt{3}}{6}\right), \quad F = \left(\sqrt{3}\frac{16+\sqrt{251}}{120}, \frac{16+\sqrt{251}}{40}, -\sqrt{3}\frac{7+\sqrt{1004}}{60}\right).$$

We have $AF=EF=1.446...$ and $DF=2.056...$. It is easy to check that the conditions of the Theorem are satisfied (Fig. 2).

Omitting the points A and E we obtain four points such that the distance $BD = BF = CF$ occurs three times, $BC = CD$ occurs twice and DF occurs once.

The author expresses her thanks to G. Fejes Tóth for his helpful suggestion.

REFERENCES

[1] Erdős, P., Combinatorial problems in geometry, *Math. Chronicle* **12** (1983), 35—54. *MR* **84i**: 52014.

[2] Erdős, P., On some metric and combinatorial geometric problems, *Discrete Math.* **60** (1986), 147—153. *MR* **88f**: 52011

[3] Rousseau, C. C., Problems, JCMN **38** (1985), 4150—4151.

[4] Palásti, I., On the seven points problem of P. Erdős, *Studia Sci. Math. Hungar.* **22** (1987), 447—448.

[5] Liu, A., On the "seven points problem" of P. Erdős, *Math. Chronicle* **15** (1986), 29—33. *MR* **88e**: 52014

[6] Palásti, I., Lattice-point examples for a question of Erdős, *Period. Math. Hungar.* **20** (1989), 231—235.

(Received June 15, 1986)

MTA MATEMATIKAI KUTATÓ INTÉZETE
P.O. BOX 127
H—1364 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 191—200

A NOTE ON f-DIVERGENCES

D. FELDMAN and F. ÖSTERREICHER

Summary

This note is based on the following representation of Csiszár's f-divergences

$$I_f(Q,P)-f(1)=\int_0^\infty [(Q-tP)^+(\Omega)-\max(1-t,0)]\,dD_+f(t),$$

for which a simple proof is given in Section 1. This representation clarifies and extends well-known results concerning these divergence measures, as exemplified by the bounds on I_f in Theorem 2 of Section 2. An application to the comparison of dichotomies is indicated in Section 3.

1. The representation theorem

Let $(\Omega,\mathscr{A})$ and $(\tilde{\Omega},\tilde{\mathscr{A}})$ be measurable spaces and let P, Q and $\tilde{P}$, $\tilde{Q}$ be probability measures on $(\Omega,\mathscr{A})$ and $(\tilde{\Omega},\tilde{\mathscr{A}})$, respectively. Furthermore, let $f\colon [0,\infty)\to\overline{\mathbf{R}}$ be a convex function and f^*, defined by $f^*(u)=uf\left(\frac{1}{u}\right)$, the $*$-conjugate convex function of f. The indeterminate form

$$0\cdot f\left(\frac{a}{0}\right)=\begin{cases} a\cdot f^*(0) & \text{for } a>0\\ 0 & \text{for } a=0\end{cases}$$

is, for $a>0$, defined by continuity. Then (cf. Csiszár [3] and Ali and Silvey [1])

$$I_f(Q,P)=\int f\left(\frac{q}{p}\right)p\,d\mu \qquad (=I_{f^*}(P,Q))$$

defines the f-divergence of the probability measures P and Q (as usual q and p denote the Radon—Nikodym-derivatives of Q and P with respect to a dominating σ-finite measure μ). The class of all functions f satisfying the above will be denoted by $\mathscr{F}$. The subclass of f's adjusted so that $f(1)=0$ will be denoted by $\mathscr{F}_0$.

Remark 1. The function $\tilde{f}\colon [0,\infty)\to\overline{\mathbf{R}}$ defined by $\tilde{f}(u)=f(u)+f^*(u)$ satisfies $(\tilde{f})^*=\tilde{f}$. Hence $I_{\tilde{f}}(Q,P)=I_{\tilde{f}}(P,Q)$.

Example 1. Let $x\in[0,1]$, $\Omega=\{0,1\}$, $P_\alpha=(\alpha,1-\alpha)$, $\alpha\in[0,1-x]$ and $f\in\mathscr{F}_0$. Then $I_f(P_{\alpha+x},P_\alpha)=\alpha f\left(1+\frac{x}{\alpha}\right)+(1-\alpha)f\left(1-\frac{x}{1-\alpha}\right)$.

1980 *Mathematics Subject Classification* (1985 Revision). Primary 94A17; Secondary 62B20.
Key words and phrases. f-divergences, comparison of dichotomies.

EXAMPLE 2. Let $x\in[0,1]$, $\Omega=\{0,1,2\}$, $\hat{P}_x=(x,1-x,0)$, $\hat{Q}_x=(0,1-x,x)$ and $f\in\mathscr{F}_0$. Then $I_f(\hat{P}_x,\hat{Q}_x)=x\tilde{f}(0)$.

EXAMPLE 3. Elementary divergences. Let

$$f_t(u)=\max(u-t,0)=u-\min(u,t)\ t\geqq 0.$$

Then the associated f-divergence is

$$I_{f_t}(Q,P)=(Q-tP)^+(\Omega)=1-b(Q,tP).$$

Here $(Q-tP)^+(\Omega)$ is the total mass of the positive part of the signed measure $Q-tP$ and $b(Q,tP)=\int\min(q,tp)\,d\mu$ is the $1+t$ — multiple of the minimum Bayes risk for the testing problem (P,Q) with respect to the prior distribution $\left(\frac{t}{1+t},\frac{1}{1+t}\right)$. Since every f-divergence is a mixture of differences

$$I_{f_t}(Q,P)-I_{f_t}(P,P)=(Q-tP)^+(\Omega)-\max(1-t,0)=\min(1,t)-b(Q,tP)$$

the family of divergences, I_{f_t}, $t\geqq 0$, may be called the family of elementary divergences.

THEOREM 1. Representation theorem. *Let $f\in\mathscr{F}_0$ and let D_+f denote the right-hand-side-derivative of the convex function f. Then*

$$I_f(Q,P)=\int[\min(1,t)-b(Q,tP)]\,dD_+f(t).$$

PROOF. An integral-geometric proof for the closely related representation (2) has been given in [13]. However, in view of the following obvious representation of a function $f\in\mathscr{F}$, satisfying $f(0)$, $D_+f(0)\in\mathbf{R}$,

$$f(u)=f(0)+uD_+f(0)+\int_0^\infty\max(u-t,0)\,dD_+f(t)$$

a much simpler analytic proof is available. Replacing u in the difference

$$f(u)-f(1)=(u-1)D_+f(0)+\int_0^\infty[\max(u-t,0)-\max(1-t,0)]\,dD_+f(t)$$

by $\frac{q}{p}$, integrating with respect to P and applying Fubini's Theorem give the result under the above assumptions. The result is easily extended to the general case using the sequence of functions $f_k(u)\uparrow f(u)$, $u\in[0,\infty)$ defined by

$$f_k(u)=\begin{cases}f\left(\frac{1}{k}\right)+\left(u-\frac{1}{k}\right)D_+f\left(\frac{1}{k}\right) & u<\frac{1}{k}\\ f(u) & u\geqq\frac{1}{k}.\end{cases}\qquad\square$$

Nevertheless it is useful to stress the geometric aspects. Hence let us recall the basic notion of the risk set of a testing problem and some of its properties. The convex set

$$R(P,Q)=\operatorname{co}\{(P(A),Q(A^c))\colon A\in\mathscr{A},\ P(A)+Q(A^c)\leqq 1\}^*$$

is called the risk set of the testing problem (dichotomy) (P, Q). Let us denote $D=\operatorname{co}\{(0,1),(1,0)\}$, $\Delta=\operatorname{co}\{(0,1),(1,0),(0,0)\}$. Then obviously:

$$D\subset R(P,Q)\subset\Delta,$$

with equality iff $P=Q$ and $P\perp Q$, respectively. Note finally that owing to

$$tP(A)+Q(A^c)\geqq b(Q,tP)$$

the risk set $R(P,Q)$ of a testing problem is determined by its family of supporting lines from below

$$\beta=b(Q,tP)-t\alpha,\qquad t\geqq 0.$$

Hence, a consequence of the representation theorem is that an f-divergence $I_f(Q,P)$ depends on the testing problem (P,Q) only via its risk set $R(P,Q)$: In fact, the following holds.

REMARK 2. Let (P,Q) and $(\bar P,\bar Q)$ be two testing problems. Then

$$R(P,Q)\supset R(\bar P,\bar Q)\quad\text{implies}\quad I_f(Q,P)\geqq I_f(\bar Q,\bar P).$$

If in addition f is strictly convex and $I_f(\bar Q,\bar P)<\infty$ then equality holds for the f-divergences iff it holds for the risk sets. This is — in essence — an immediate consequence of the relation:

$$R(P,Q)\supset R(\bar P,\bar Q)\Leftrightarrow b(Q,tP)\leqq b(\bar Q,t\bar P)\quad\forall t\geqq 0$$

and the obvious generalization of the representation theorem

$$I_f(Q,P)-I_f(\bar Q,\bar P)=\int_0^\infty[b(\bar Q,t\bar P)-b(Q,tP)]\,dD_+f(t)\tag{1}$$

for $f\in\mathscr{F}$ and $(\bar P,\bar Q)$ such that $I_f(\bar Q,\bar P)<\infty$.

Corresponding to the two extremes D and Δ for the risk set there are two extremes for $I_f(Q,P)$, the measure of information contained in (P,Q):

$I_f(Q,P)=0$ for $Q=P$, i.e. for the least informative dichotomy and

$I_f(Q,P)=\tilde f(0)$ for $Q\perp P$, i.e. for the most informative dichotomy.

In Section 2 we are going to deal with $I_f(Q,P)=I_f(Q,P)-I_f(P,P)$, i.e. with the comparison of (P,Q) with the least informative dichotomy. For the comparison of (P,Q) with the most informative dichotomy we need to assume $\tilde f(0)<\infty$. In this case the corresponding measure

$$\tilde f(0)-I_f(Q,P)=\int_0^\infty b(Q,tP)\,dD_+f(t)\tag{2}$$

* $\operatorname{co}(S)$ denotes the convex hull of the set S.

is most appropriately stated in terms of the concave function

$$g(u)=f(0)+uf^*(0)-f(u) \quad \text{as}$$

(3)
$$\bar{I}_g(Q,P)=\int g\left(\frac{q}{p}\right)p\,d\mu=\int_0^\infty b(Q,tP)\,dD_+g(t).$$

2. Bounds on I_f and their properties

Consider the function $c_f\colon [0,1]\to\overline{\mathbf{R}}$ defined by

(4)
$$c_f(x)=\min_{0\le\alpha\le 1-x}\{I_f(P_{\alpha+x},P_\alpha)\}$$

where P_α is as in Example 1.

The function c_f is to be used to obtain bounds on the f-divergence $I_f(Q,P)$. These bounds are derived as a corollary of Theorem 1 (Remark 2).

Corollary 1. *For $f\in\mathscr{F}_0$ and any testing problem (P,Q) such that $(Q-P)^+(\Omega)=x$ the following inequalities hold and are best possible:*

$$c_f(x)\le I_f(Q,P)\le xc_f(1).$$

Proof. Let $\hat{P}_x$ and $\hat{Q}_x$ be as in Example 2. Then

$$\{(P,Q)\colon (Q-P)^+(\Omega)=x\}=$$
$$=\{(P,Q)\colon R(P,Q)\subset R(\hat{P}_x,\hat{Q}_x) \text{ and } \exists\alpha\in[0,1-x]\colon R(P_{\alpha+x},P_\alpha)\subset R(P,Q)\}.$$

Then Remark 2, Example 2 and the definition of c_f yield

$$c_f(x)\le I_f(P_{\alpha+x},P_\alpha)\le I_f(Q,P)\le I_f(\hat{Q},\hat{P})=c_f(1)\cdot x. \quad \square$$

The following theorem gives important properties of the function c_f.

Theorem 2. *For $f\in\mathscr{F}_0$ the function c_f satisfies*

(a)
$$c_f(x)\ge\tilde{f}(1-x)\ge 0 \qquad \forall x\in[0,1]$$

with strict second in equality for $x\in(0,1]$ iff f is strictly convex at $u=1$. Furthermore $c_f(0)=0$ and $c_f(1)=\tilde{f}(0)$.

(b) *c_f is convex and continuous on* $[0,1]$.

(c) *c_f is increasing (strictly increasing if f is strictly convex at $u=1$).*

(d) $c_{f^*}=c_f\le\frac{1}{2}c_{\tilde{f}}$. *If, in addition, $f^*=f$ then*

$$c_f(x)=\frac{1}{2}c_{\tilde{f}}(x)=(1+x)f\left(\frac{1-x}{1+x}\right).$$

PROOF. (a) follows from

$$\max\{b(P_{\alpha+x}, tP_\alpha), \alpha\in[0, 1-x]\} = \begin{cases} b(P_x, tP_0) & \text{for } 0\leq t\leq 1 \\ b(P_1, tP_{1-x}) & \text{for } 1<t \end{cases}$$

and the representation theorem.

(b) The convexity of c_f is based on the fact that the function h: $[0, \infty)^2\to\overline{\mathbf{R}}$, defined by $h(x, y)=xf\left(\frac{y}{x}\right)$, is a convex function of the vector (x, y): Let $x_i\in[0, 1]$ and $\alpha_i\in[0, 1-x_i]$ such that

$$c_f(x_i) = I_f(P_{\alpha_i+x_i}, P_{\alpha_i}) \qquad i\in\{1, 2\}.$$

Then

$$\gamma I_f(P_{\alpha_1+x_1}, P_{\alpha_1})+(1-\gamma)\, I_f(P_{\alpha_2+y_2}, P_{\alpha_2}) \geq I_f(P_{\gamma\alpha_1+(1-\gamma)\alpha_2+\gamma x_1+(1-\gamma)x_2}, P_{\gamma\alpha_1+(1-\gamma)\alpha_2}) \geq$$
$$\geq c_f(\gamma x_1+(1-\gamma)\, x_2).$$

(c) follows immediately from (a) and (b).

(d) finally, is an easy consequence of

$$I_f(Q, P) = I_{f^*}(P, Q) \quad\text{and}\quad c_f(x)\leq I_f(P_{\alpha(x)+x}, P_{\alpha(x)}) \quad\text{with}\quad \alpha(x)=\frac{1-x}{2}. \qquad \square$$

We now give further examples of f-divergences.

EXAMPLE 4.

$$f(u) = |u-1|^s, \qquad s\geq 1$$

$$c_f(x) = \begin{cases} (2x)^s & x\leq \frac{1}{2} \\ x\left(1+\left(\frac{x}{1-x}\right)^{s-1}\right) & x>\frac{1}{2}. \end{cases}$$

The f-divergences associated with this class of f's contain (for $s=1$ and $s=2$) the total-variation distance and the χ^2-divergence.

EXAMPLE 5.

$$f(u) = \frac{1}{2}(1+u)-\sqrt{u} = \frac{1}{2}[1-\sqrt{u}]^2,$$

$$c_f(x) = 1-\sqrt{1-x^2}.$$

The f-divergence for this f is the square of the so-called Hellinger-distance. The inequality of Corollary 1 applied to this c_f is well-known and has been derived, e.g., in Kraft [9], Lemma 1.

EXAMPLE 6.

$$f(u) = \frac{1}{2}(1+u)-2\frac{u}{1+u} = \frac{1}{2}\frac{[1-u]^2}{1+u}$$

$$c_f(x) = x^2.$$

The f-divergence associated with this f has been investigated, e.g., in Vincze [16]. Its squareroot is another distance.

EXAMPLE 7.

$$f(u) = \sqrt{1+u^2}-\sqrt{2} \qquad \text{(cf. [14])}$$

$$c_f(x) = \sqrt{2}\,(\sqrt{1+x^2}-1).$$

We conjecture that $\sqrt{I_f}$ is also a distance for this case.[1]

EXAMPLE 8.

$$f(u) = \begin{cases} \frac{1}{2}(u-1)^2 & \text{for} \quad u\in[1-\theta,\, 1+\theta] \\ \theta\,|u-1|-\frac{\theta^2}{2} & \text{otherwise} \end{cases} \qquad \theta\in(0,\, 1]$$

$$c_f(x) = \begin{cases} 2x^2 & \text{for} \quad x\in[0,\, \theta/2] \\ 2\theta x-\frac{\theta^2}{2} & \text{for} \quad x\in(\theta/2,\, 1]. \end{cases}$$

The f-divergences for this class of f's will be used in the proof of Theorem 3.

Since only in exceptional cases can $c_f(x)$ be given explicitly, reasonable lower bounds for $c_f(x)$ are desirable. One is given in Theorem 2, part (a). Lower bounds of the form kx^2 has been extensively studied for the special case $f(u)=u \ln u$ (see below). We will consider such bounds for general f in the subclass $\mathscr{F}_2$ which contains the most commonly considered f's:

$\mathscr{F}_2=\{f\in\mathscr{F}_0$, for which the second derivative $f^{(2)}$ exists and is positive in some neighbourhood of $1\}$.

THEOREM 3. *Let* $f\in\mathscr{F}_2$ *and let* $k_f=\max\{k\geqq 0 : kx^2\leqq c_f(x)\ \forall x\in[0,\, 1]\}$. *Then*

$$0 < \max\{f_1(\theta)\, h(\theta),\, \theta\in[0,\, \infty)\} \leqq k_f \leqq \min\{\tilde{f}(0),\, 2f^{(2)}(1)\},$$

where

$$f_1(\theta) = \min\{f^{(2)}(u),\, u\in[\max(0,\, 1-\theta),\, 1+\theta]\}$$

and

$$h(\theta) = \begin{cases} 2\theta-\frac{\theta^2}{2} & \theta\in[0,\, 1] \\ \min\left(\frac{1}{2}+\theta,\, 2\right) & \theta\in(1,\, \infty). \end{cases}$$

PROOF. $k_f\leqq\tilde{f}(0)$ follows from the definition of k_f by setting $x=1$. For the second upper bound we have to study the behaviour of $c_f(x)$ at $x=0$. Let $\theta\in[0,\, 1]$ and $f_2(\theta)=\max\{f^{(2)}(u),\, u\in[1-\theta,\, 1+\theta]\}$. Then

$$(u-1)^2 f_2(\theta)/2 \geqq f(u)-(u-1)f^{(1)}(1) \geqq (u-1)^2 f_1(\theta)/2 \quad \forall u\in[0,\, \theta/2].$$

[1] Meanwhile this conjecture has been verified. Cf. Kafka, P., Österreicher, F. and Vincze, I., On powers of f-divergences defining a distance (submitted to *Studia Sci. Math. Hungarica*).

Hence by Example 8

$$2f_2(\theta) \geqq c_f(x)/x^2 \geqq 2f_1(\theta) \quad \forall x \in [0, \theta/2]$$

and therefore, $\lim_{x \downarrow 0} c_f(x)/x^2 = 2f^{(2)}(1) \geqq k_f$. In addition, Example 8 provides (at $x=1$) the lower bound

$$\max\left\{f_1(\theta)\left(2\theta - \frac{\theta^2}{2}\right),\ \theta \in [0, 1]\right\}.$$

The rest follows by considering the straightforward modification of Example 8 for $\theta > 1$. □

Theorem 2, parts (a) and (d), give some further information concerning k_f. Finally we state another sufficient condition for $k_f < 2f^{(2)}(1)$.

Remark 3. Let $f \in \mathscr{F}_2$ have a continuous fourth derivative. Then

$$-\frac{4}{3}\frac{[f^{(3)}(1)]^2}{f^{(2)}(1)} + f^{(4)}(1) < 0 \quad \text{implies} \quad k_f < 2f^{(2)}(1).$$

A simple consequence of Corollary 1 for $f \in \mathscr{F}_2$ is

$$\|Q-P\| \leqq \sqrt{c I_f(Q, P)}$$

where $c \in (0, \infty)$ and $\|Q-P\| = 2(Q-P)^+(\Omega)$, the total-variation-norm of P and Q. This inequality has already been proved for general $f \in \mathscr{F}_2$ by Csiszár [6]. The best possible c is, of course, $c = 4/k_f$.

For the most prominent example: $f(u) = u \ln u$, which gives Kullback—Leibler's I-divergence $I(Q\|P)$, there is a challenging history concerning the optimal constant c: Pinsker (1960) proved $c < \infty$; Csiszár (1966): $c < 16$; McKean (1966): $c < 4e$; Csiszár (1967), Kullback (1967), Kemperman (1968): $c = 2$. (Our general conclusions concerning k_f applied to this case yields $c \in [2, 4)$.)

Example 9.

$$f(u) = u \ln u$$

$$c_f(x) \geqq \max\{2x^2, -x \ln x\}.$$

$k_f = 2f^{(2)}(1)$ holds also for the following extension of Example 5.

Example 5 (extended).

$$f(u) = \frac{1}{2}(1+u) - u^\alpha, \quad \alpha \in (0, 1)$$

$$c_f(x) \geqq \max\{2\alpha(1-\alpha)x^2, 2 - x - (1-x)^\alpha - (1-x)^{1-\alpha}\}.$$

(Our conclusions concerning k_f yield $\alpha(1-\alpha) \leqq k_f \leqq 2\alpha(1-\alpha)$.)

3. Application to comparison of dichotomies

A concave function g: $[0, \infty)\to\mathbf{R}$ satisfying $g(0)=g^*(0)=0$ can be associated with a decision problem of the following form:

Let $\Theta = \{0, 1\}$ be the parameter space, $A = [0, 1]$ the action space and L: $\Theta\times A\to[0, \infty)$ a loss function satisfying $L(0, 0)=L(1, 1)=0$. Furthermore, let $(\xi, 1-\xi)$ be a prior distribution on Θ and $L_\xi(a)=\xi L(0, a)+(1-\xi)L(1, a)$ the Bayes risk of the action a when the prior distribution is $(\xi, 1-\xi)$. Finally let

$$U(1-\xi) = \min_{0\leq a\leq 1} L_\xi(a)$$

be the minimal Bayes risk. Then we associate g with the concave function U: $[0, 1]\to$ $\to[0, \infty)$ satisfying $U(0)=U(1)=0$ via

$$g(u) = (1+u)U\left(\frac{u}{1+u}\right).$$

EXAMPLE 5 (continued). Let $\alpha\in(0, 1)$ and $L(\theta, a)=|1-\alpha-\theta|\left(\frac{a}{1-a}\right)^{\alpha-\theta}$, $\theta\in\{0, 1\}$. Then $U(1-\xi)=L_\xi(1-\xi)=(1-\xi)^\alpha\xi^{1-\alpha}$ and $g(u)=u^\alpha$.

EXAMPLE 6 (continued). Let $L(\theta, a)=2(\theta-a)^2$, $\theta\in\{0, 1\}$. Then $U(1-\xi)=$ $=L_\xi(1-\xi)=2\xi(1-\xi)$ and $g(u)=2\frac{u}{1+u}$. $U/4$ was used by Le Cam [11] to define a "testing affinity" for a pair of probability measures.

Let x be an observation of a random variable distributed according to one of the probability distribution of the dichotomy (P, Q). Furthermore, let $(\xi_x, 1-\xi_x)$, where

$$\xi_x = \frac{\xi p(x)}{\xi p(x)+(1-\xi)\,q(x)},$$

be the posterior distribution on $\Theta=\{0, 1\}$. Then the expected posterior risk, for any specified U, is

$$U_\xi(Q, P) = \int U(1-\xi_x)\,(\xi p(x)+(1-\xi)\,q(x))\,d\mu(x),$$

and can be used to induce a comparison of dichotomies as follows: (P, Q) is more informative than $(\bar{P}, \bar{Q})$ with respect to U if

$$U_\xi(Q, P) \leq U_\xi(\bar{Q}, \bar{P}) \quad \forall\xi\in(0, 1).$$

This comparison can be equivalently expressed in terms of the family of the concave functions $g_s(u)=sg\left(\frac{u}{s}\right)$ by

$$I_{g_s}(Q, P) \leq I_{g_s}(\bar{Q}, \bar{P}) \quad \forall s\in(0, \infty),$$

which can in turn be generalized to the following definition: (P, Q) is called more informative than $(\tilde{P}, \tilde{Q})$ with respect to f, in short $(P, Q) >_f (\tilde{P}, \tilde{Q})$, if

$$I_{f_s}(Q, P) \geqq I_{f_s}(\tilde{Q}, \tilde{P}) \quad \forall s \in (0, \infty),$$

where $f_s(u) = sf\left(\frac{u}{s}\right)$ and $f \in \mathcal{F}$.

Among the orderings of this type is the one corresponding to $f(u) = \max(u-1, 0)$ (i.e. the one induced by the elementary divergences). For this ordering we obtain

$$(P, Q) >_f (\tilde{P}, \tilde{Q}) \quad \text{iff} \quad R(P, Q) \supset R(\tilde{P}, \tilde{Q}),$$

which was first shown by Blackwell [2]. It is the strongest ordering and hence yields the smallest class of comparable dichotomies. It was shown in [7] that on the other extreme the following U's, and only these, yield a total ordering of dichotomies:

$$U(1-\xi) = c(1-\xi)^\alpha \xi^{1-\alpha} \qquad \alpha \in (0, 1), \; c > 0.$$

In view of (1) those $f \in \mathcal{F}$ which have a (continuous) second derivative $f^{(2)}$ on $(0, \infty)$, satisfying $f_s^{(2)}(u) = c(s) f^{(2)}(u)$ for some function $c(s)$, yield a total ordering of dichotomies. For these f's $f^{(2)}(u) = cu^\mu$, $\mu \in \mathbf{R}$, $c \geqq 0$ and hence

$$f(u) = a + bu + \frac{c}{(\mu+1)(\mu+2)} u^{\mu+2} \qquad \mu \in \mathbf{R} \setminus \{-1, -2\} \tag{0}$$

$$f(u) = a + bu + c\, u \ln u \tag{1}$$

$$f(u) = a + bu - c \ln u \qquad a, b \in \mathbf{R}. \tag{2}$$

Note that the U's yielding a total ordering correspond to f's with $\mu \in (-2, -1)$. The classification of those U suggests the conjecture that the only $f \in \mathcal{F}$ yielding a total ordering are the above.

REFERENCES

[1] Ali, S. M. and Silvey, S. D., A general class of coefficients of divergence of one distribution from another, *J. Roy. Statist. Soc. Ser. B* **28** (1966), 131—142. *MR* **33** # 4963.

[2] Blackwell, D., Comparison of experiments, *Proc. Second Berkeley Symp. Math. Statist. Prob.* 1950, University of California Press, Berkeley and Los Angeles, 1951, 93—102. *MR* **13**—667.

[3] Csiszár, I., Eine informationstheoretische Ungleichung und ihre Anwendung auf den Beweis der Ergodizität von Markoffschen Ketten, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **8** (1963), 85—108. *MR* **29** # 1671.

[4] Csiszár, I., A note on Jensen's inequality, *Studia Sci. Math. Hungar.* **1** (1966), 185—188. *MR* **35** # 5563.

[5] Csiszár, I., Information-type measures of difference of probability distributions and indirect observations, *Studia Sci. Math. Hungar.* **2** (1967), 299—318. *MR* **36** # 2428.

[6] Csiszár, I., On topological properties of *f*-divergences, *Studia Sci. Math. Hungar.* **2** (1967), 329—339. *MR* **36** # 2429.

[7] Feldman, D., Some properties of Bayesian orderings of experiments, *Ann. Math. Statist.* **43** (1972), 1428—1440. *MR* **49** # 8151.

[8] Kemperman, J. H. B., On the optimum rate of transmitting information, *Probability and information theory* (Proc. Internat. Sympos., McMaster Univ., Hamilton, Ont., 1968), Lecture Notes in Mathematics, Vol. 89, Springer-Verlag, Berlin, 1969, 126—169. *MR* **42** # 5720.

[9] KRAFT, C., Some conditions for consistency and uniform consistency of statistical procedures, *Univ. California Publ. Statist.* **2** (1955), 125—141. *MR* **17**—505.

[10] KULLBACK, S., A lower bound for discrimination information in terms of variation, *IEEE Trans. Information Theory*, IT—**13** (1967), 126—127.

[11] LE CAM, L., On the asymptotic normality of estimates, *Proceedings of the symposium to honour Jerzy Neyman* (Warsaw, 1974), Pańistw. Wydawn. Nauk, Warsaw, 1977. *MR* **57** # 10882.

[12] MCKEAN, H. P., Speed of approach to equilibrium for Kac's caricature of a Maxwellian gas, *Archiv Rat. Mech. Anal.* **21** (1966), 343—367.

[13] ÖSTERREICHER, F. and FELDMAN, D., Divergenzen von Wahrscheinlichkeitsverteilungen—integralgeometrisch betrachtet, *Acta Math. Acad. Sci. Hungar.* **37** (1981), 329—337. *MR* **83b**: 62043.

[14] ÖSTERREICHER, F., The construction of least favourable distributions is traceable to a minimal perimeter problem, *Studia Sci. Math. Hungar.* **17** (1982), 341—351. *MR* **85k**: 62009.

[15] PINSKER, M. S., Information and information stability of random variables and processes, Holden-Day, San Francisco, 1964 (Russian original 1960).

[16] VINCZE, I., On the concept and measure of information contained in an observation, *Contributions to probability* (ed. by J. Gani and V. K. Rohatgi), Academic Press, New York, 1981, 207—214. *MR* **83b**: 62007.

(Received June 16, 1986)

DEPARTMENT OF STATISTICS AND PROBABILITY
MICHIGAN STATE UNIVERSITY
EAST LANSING, MI 48824—1027
U.S.A.

INSTITUT FÜR MATHEMATIK
UNIVERSITÄT SALZBURG
HELLBRUNNERSTRASSE 34
A—5020 SALZBURG
AUSTRIA

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 201—211

ON A SECOND ORDER ALGEBRAIC DIFFERENTIAL EQUATION

TAMÁS FÉNYES

Introduction

In the paper [1] we have discussed the algebraic differential equation

$$D^2(x)-\left(\frac{D(f)}{f}+2m\right)D(x)+\left(m\frac{D(f)}{f}+m^2-b^2f^2\right)x = y \tag{1}$$

defined in the discrete Mikusiński operator field M based on the Cauchy product. In (1) $x\in M$ is the unknown solution, if it exists. D is the symbol of the algebraic derivative, f and y are arbitrary functions defined for $n=0, 1, 2, \ldots,$ with the restriction $f(0)=0$, m and $b\neq 0$ are arbitrarily given real numbers.

In the present paper we shall give an operational treatment of (1) in the case when the discrete operator field M is based on the so-called number-theoretical Dirichlet product

$$\{a(n)\}\{b(n)\} = \left\{\sum_{v/n} a(v)b\left(\frac{n}{v}\right)\right\}$$

of functions defined on the set of natural numbers.

The operational treatment of (1) is motivated by the fact, that it leads to an interesting operational Alternative Theory.

For the theory and applications of the discrete operational calculus see [2], [3], [4].

In Chapter 1 we briefly summarize the results of the paper [2], giving some generalizations of them, Chapter 2 contains the operational theory of the differential equation (1).

In what follows **Z** will denote the set of natural numbers.

§ 1. Discrete Mikusiński operators based on the Dirichlet product

Let $a=\{a(n)\}$ be an arbitrary real-valued function defined on Z. The symbol $a(n)$ denotes the value of this function for arbitrary fixed n.

Let E denote the set of the discrete functions. If we introduce in E the following two operations

(i) $a+b := \{a(n)\}+\{b(n)\} = \{a(n)+b(n)\}$, addition

(ii) $ab := \{a(n)\}\{b(n)\} = \left\{\sum_{v/n} a(v)b\left(\frac{n}{v}\right)\right\}$, multiplication,

---

1980 *Mathematics Subject Classification* (1985 Revision). Primary 44A40.
Key words and phrases. Mikusiński's operator calculus, operational differential equations.

then E becomes a commutative ring without zero divisor and can be extended to a quotient field.

This is called the discrete Mikusiński operator field and is denoted by M. The elements of M are called M-operators.

The definition and properties of the "discrete" Dirac-function

We define the discrete Dirac-function by

$$\delta(N) = \{\delta(n, N)\},$$

where

$$\delta(n, N) = \begin{cases} 0, & \text{for} \quad n \neq N, \\ 1, & \text{for} \quad n = N. \end{cases} \qquad N \in \mathbf{Z}.$$

For later purposes we enumerate some properties of the Dirac function.

PROPERTY 1.

$$\delta(N)\{a(n)\} = \{b(n)\},$$

(1.1)
$$b(n) = \begin{cases} a\left(\dfrac{n}{N}\right), & \text{for} \quad N/n, \\ 0 & \text{otherwise.} \end{cases}$$

(1.2)
$$\delta(N_1)\,\delta(N_2) = \delta(N_1 N_2), \qquad N_1, N_2 \in \mathbf{Z}$$

PROPERTY 2.

(1.3)
$$x = \frac{\{a(n)\}}{\delta(N)} \in E, \qquad N \in \mathbf{Z}$$

holds if and only if

(1.4)
$$a(n) = 0$$

for those values of n for which N is not a divisor of n. If (1.4) holds, then

(1.5)
$$x = \{a(n, N)\}.$$

The field K of the real numbers can be embedded isomorphically into the operator field M. The common unit element of K, E, M is the function $\delta(1)$ and we write

$$\delta(1) = 1.$$

Moreover,

$$c\delta(1) = c, \qquad c\{a(n)\} = \{ca(n)\}$$

for every $c \in K$ and every $a \in E$. Every operator of the form

$$x = \frac{\{a(n)\}}{\{b(n)\}}$$

is a function if $b(1) \neq 0$.

The operator function $\delta(\varepsilon)$.

For arbitrary rational number $\varepsilon=\frac{N_1}{N_2}$ we define

$$\delta(\varepsilon)=\frac{\delta(N_1)}{\delta(N_2)}. \tag{1.6}$$

From this definition it follows that for $\varepsilon\in N$ we have

$$\delta(\varepsilon)=\delta(N)=\{\delta(n, N)\}.$$

If

$$\frac{N_1}{N_2}=\frac{N_3}{N_4},$$

then

$$\delta\left(\frac{N_1}{N_2}\right)=\delta\left(\frac{N_3}{N_4}\right)$$

holds.

PROPERTY 3. Let α, β be arbitrary positive rational numbers, then

$$\delta(\alpha)\delta(\beta)=\delta(\alpha\beta) \tag{1.7}$$

and it is easily seen that

$$\delta\left(\frac{1}{\alpha}\right)=\frac{1}{\delta(\alpha)} \tag{1.8}$$

is also true.

The definition of the ring E^*.

Let $E^*\subset M$ be the subset of M whose elements are of the form

$$x=\frac{a}{\delta(N)} \qquad N\in\mathbf{Z},\ a\in E. \tag{1.9}$$

E^* is a ring and, by choosing $N=1$, we have

$$E\subset E^*$$

PROPERTY 4. Obviously,

$$x=\frac{a}{\delta(\varepsilon)}\in E^*, \qquad \varepsilon=\frac{N_1}{N_2}$$

(N_1, N_2 are relatively primes).

Moreover, $x\in E$ if and only if

$$a(n)=0$$

for those values of n for which N_1 is not a divisor of n. If the condition is satisfied, we have

$$x(n)=\begin{cases} a\left(\dfrac{nN_1}{N_2}\right), & \text{for} \quad N_2|n, \\ 0 & \text{otherwise.}\end{cases}$$

Definition of the convergence in the ring E.

Let $a_k \in E$, $(k=1, 2, \ldots)$ be an infinite sequence of functions. By definition

$$\lim_{k\to\infty} \{a_k(n)\} = \{a(n)\} \tag{1.10}$$

if for every fixed n

$$\lim_{k\to\infty} a_k(n) = a(n)$$

(see [5]). This convergence can be extended to infinite series of functions as usual. Let

$$f(z) = \sum_{k=0}^{\infty} \beta_k z^k \qquad \beta_k \in K$$

be an arbitrary entire function of the complex variable z. Then

$$f(a) = \sum_{k=0}^{\infty} \beta_k \{a(n)\}^k, \quad a \in E,\ a^0 = 1 \tag{1.11}$$

holds in the sense of the convergence defined above. We have

$$e^a = \sum_{k=0}^{\infty} \frac{a^k}{k!} \qquad a \in E$$

having the property

$$e^a e^b = e^{a+b}, \qquad a, b \in E,$$

moreover, if we write

$$e^a = \{e_a(n)\},$$

so

$$e_a(1) = e^{a(1)} \tag{1.12}$$

holds.

Moreover let

$$\sum \gamma_k z^k \qquad \gamma_k \in K$$

be an arbitrary *formal* infinite series and let $a \in E$ an arbitrary function with $a(1)=0$. Then

$$\sum_{k=0}^{\infty} \gamma_k a^k$$

also converges in the sense of convergence defined above.

The algebraic derivation and integration (see also [4]).

For the sake of easy reading we recapitulate some definitions and facts of the algebraic derivation and integration.

$$D(a) = \{-\log n \cdot a(n)\}, \qquad a \in E \tag{1.13}$$

$$D\left(\frac{a}{b}\right) = \frac{bD(a) - aD(b)}{b^2}, \qquad a, b \in E, \quad \frac{a}{b} \in M. \tag{1.13'}$$

PROPERTY 5.

$$D\left[\frac{a}{\delta(\varepsilon)}\right] = \frac{\left\{-\log\frac{n}{\alpha}\cdot a(n)\right\}}{\delta(\varepsilon)} \in E^* \qquad a \in E,\ \varepsilon = \frac{N_1}{N_2}. \tag{1.14}$$

$$D[\delta(\varepsilon)] = -\log\varepsilon\cdot\delta(\varepsilon). \tag{1.14'}$$

PROPERTY 6.

$$D(e^a) = D(a)\,e^a, \qquad a \in E. \tag{1.15}$$

If for a given $x \in M$ there exists a $y \in M$ such that

$$D(y) = x,$$

we say that x is algebraic integrable and we write

$$y = \int x.$$

PROPERTY 7. If $x \in M$ and

$$D(x) = 0,$$

then x is an arbitrary number.

Two algebraic integrals of an operator may differ only by an arbitrary number.

The algebraic differentiation and integration is a linear operation over the field of the real (complex) numbers.

PROPERTY 8. The operator

$$x = \frac{a}{\delta(\varepsilon)}, \qquad a \in E,\ \varepsilon = \frac{N_1}{N_2} \tag{1.16}$$

is algebraic integrable in M if and only if either

$$\varepsilon \neq N, \quad N \in Z,$$

or $\varepsilon \in N$ and $a(N)=0$ holds true. Every algebraic integral of (1.16) belonging to E^* is given by

$$\int \frac{a}{\delta(\varepsilon)} = \frac{\left\{-\dfrac{a(n)}{\log\dfrac{n}{\varepsilon}}\right\}}{\delta(\varepsilon)} + c, \qquad c \in K \tag{1.17}$$

where in the case of $\varepsilon = N$ the symbol

$$\frac{a(N)}{\log\dfrac{N}{N}}$$

denotes an arbitrary real (complex) number. We shall choose this to be null.

For $\varepsilon=1$ we have that a is integrable if and only if $a(1)=0$, and

$$\int a = \left\{-\frac{a(n)}{\log n}\right\}+c, \quad c\in K.$$

§ 2. On the second order differential equation (1)

Let us consider the algebraic differential equation

$$D^2(x)-\left(\frac{D(f)}{f}+2m\right)D(x)+\left(m\,\frac{D(f)}{f}+m^2-b^2f^2\right)x = y \tag{2.1}$$

where $y\in E$, $f\in E$, $m\in K$, $b\in K$ are given arbitrarily, with the restrictions $f\neq 0$, $f(1)=0$, $b\neq 0$.

First we discuss the homogeneous equation $(y=0)$. We prove the following

THEOREM 1. *Let* $y=0$, $f(1)=0$, $b\neq 0$. *The differential equation* (2.1) *has a nontrivial solution in the operator field* M *if and only if*

$$\alpha = e^{-m}$$

is rational. Moreover, in this case there exist two linearly independent solutions belonging to the ring E^*.

PROOF. *Sufficiency.* Let $\alpha=e^{-m}$ be rational and let us apply the substitution

$$x = \delta(\gamma)^u \quad u\in M,\ \gamma = \frac{R_1}{R_2}, \quad R_1, R_2\in \mathbf{Z}.$$

Taking into account (1.14) after some calculation we have

$$D^2(u)-\left[2\log\gamma+2m+\frac{D(f)}{f}\right]D(u)+ \tag{2.2}$$

$$+\left[\log^2\gamma+2m\log\gamma+m^2+(m+\log\gamma)\,\frac{D(f)}{f}-b^2f^2\right]u = 0.$$

By choosing $\gamma=\alpha=e^{-m}$ we have

$$D^2(u)-\frac{D(f)}{f}D(u)-b^2f^2u = 0. \tag{2.3}$$

One can easily see that the linearly independent solutions of (2.3) are

$$u_1 = e^{b\int f}, \quad u_2 = e^{-b\int f}.$$

From Property 8 of Chapter 1 follows that $\int f$ exists, since $f(0)=0$. So the general solution of (2.1) is of the form

$$x = c_1x_1+c_2x_2 \qquad c_1, c_2\in K \tag{2.4}$$

where

$$x_1 = \delta(\alpha)\, e^{b\int f}, \qquad \alpha \in e^{-m}$$

(2.5)
$$x_2 = \delta(\alpha)\, e^{-b\int f}.$$

One can see that $x_1, x_2 \in E^*$.

Necessity. In the paper [3] we have shown that for every M-operator

$$Y = \frac{p}{q}, \qquad p, q \in E$$

$$r = \frac{D(Y)}{Y} \in E$$

holds, where

$$r(1) = \log \frac{N_q}{N_p}.$$

Here the numbers N_p, N_q are the smallest integers for which the functions p and q do not vanish, respectively. Since

$$D\left[\frac{D(Y)}{Y}\right] = \frac{D^2(Y)}{Y} - \left[\frac{D(Y)}{Y}\right]^2,$$

it can easily be seen that the operator $\dfrac{D^2(Y)}{Y}$ is also a function having for $n=1$ the value $\log^2 \dfrac{N_q}{N_p}$.

Now let Y be a nontrivial solution of (2.1). Let us write (2.1) for $y=0$ in the form

(2.5)
$$\frac{D^2(Y)}{Y} - \left(\frac{D(f)}{f} + 2m\right)\frac{D(Y)}{Y} + m\frac{D(f)}{f} + m^2 - b^2 f^2 = 0.$$

Since $f(1)=0$, we obtain for $n=1$

(2.6)
$$\log^2 \frac{N_q}{N_p} - (2m - \log R) \log \frac{N_q}{N_p} - m \log R + m^2 = 0,$$

where $-\log R$ is the value of the function $\dfrac{D(f)}{f}$ for $n=1$. (2.6) is an algebraic equation of second degree for m having the roots

$$m_1 = \log \frac{N_q}{N_p}, \quad m_2 = \log \frac{RN_q}{N_p}.$$

So the numbers

$$e^{-m_1}, \ e^{-m_2}$$

are rational and the theorem is proved.

THEOREM 2. *Let* $y=0$, $f(1)=0$, $b\neq 0$ *and let*

$$\alpha = e^{-m} = \frac{N_2}{N_1}, \quad N_1, N_2 \in Z$$

where N_1, N_2 *are relatively primes. If* α *is an integer, then* (2.1) *has two linearly independent solutions in the ring* E. *If* α *is not an integer, then* (2.1) *has a nontrivial solution in* E *if and only if* f *vanishes for those values on* n *for which* N_1 *is not a divisor of* n. *Moreover, in the case of non-integer* α, (2.1) *cannot have two linearly independent solutions in the ring* E.

PROOF. Taking into account (2.5) it is easily seen that the first half of the theorem is a simple consequence of Property 1 of Chapter 1.

Let now $N_1>1$. We discuss the operator

$$x = \frac{x_1 - x_2}{2} = \delta(\alpha) \operatorname{sh} b \int f. \tag{2.7}$$

From Property 4 of Chapter 1 it follows that $x\in E$ if and only if $\operatorname{sh} b\int f$ vanishes for the values of n given in the Theorem.

However, if f vanishes for these values of n, then $b\int f$ also vanishes trivially there.

Applying the expansion

$$\operatorname{sh} b \int f = \sum_{i=0}^{\infty} \frac{b^{2i+1} f^{2i+1}}{(2i+1)!},$$

we get that $\operatorname{sh} b\int f$ also vanishes for the values n given above, since f^{2i+1} vanishes there.

Conversely, if $\operatorname{sh} b\int f$ vanishes for these values of n, then applying the expansion

$$b \int f = \operatorname{arsh}\left(\operatorname{sh} b \int f\right) = \operatorname{sh} b \int f - \frac{1}{2.3}\left(\operatorname{sh} b \int f\right)^3 + \frac{1.3}{2.4.5}\left(\operatorname{sh} b \int f\right)^5 \ldots,$$

(which converges trivially since for every fixed n the infinite series reduces to a finite one) it can be seen that $b\int f$, consequently also f vanishes for the values of n given in the theorem.

Finally, it remains to show that the homogeneous equation cannot have two linearly independent solutions in the ring E.

Obviously, the general solution of (2.1) can be written also in the form

$$x = \delta(\alpha)\left[c_1 \operatorname{sh} b \int f + c_2 \operatorname{ch} b \int f\right], \quad c_1, c_2 \in K.$$

If $x\in E$, then also

$$x\delta\left(\frac{1}{\alpha}\right) = x\delta\left(\frac{N_1}{N_2}\right) = c_1 \operatorname{sh} b \int f + c_2 \operatorname{ch} b \int f \in E.$$

By (1.12) sh $b\int f$ equals to zero for $n=1$, ch $b\int f$ equals to one for $n=1$. From Property 4 of Chapter 1 follows that $x\delta\left(\frac{1}{\alpha}\right)$ equals to zero for $n=1$.

So we have $c_2=0$, and the theorem is proved.

In the foregoing we shall deal with the inhomogeneous problem. We distinguish two cases according that $\alpha=e^{-m}$ is rational or not.

I. $$\alpha = e^{-m} = \frac{N_2}{N_1} \qquad N_1, N_2 \in \mathbf{Z}.$$

As we have seen, the homogeneous equation has now nontrivial solutions. From a previous result of the author follows, that (2.1) has a solution in the field of the M-operators if and only if the algebraic integrals

$$\int \frac{yx_1}{w}, \quad \int \frac{yx_2}{w}$$

exist, where w denotes the Wronski determinant (see [6], page 339). Moreover, a particular solution of (2.1) is of the form

(2.8) $$x_p = x_2 \int \frac{yx_1}{w} - x_1 \int \frac{yx_2}{w}.$$

Since by (2.5), (1.15)

(2.9) $$w = x_1 D(x_2) - x_2 D(x_1) = D\left(\frac{x_2}{x_1}\right) x_1^2 = -2bf\delta^2(\alpha),$$

we get the formal solution

(2.10) $$x_p = \frac{\delta(\alpha)}{2b} e^{b\int f} \int \frac{ye^{-b\int f}}{f\delta(\alpha)} - \frac{\delta(\alpha)}{2b} e^{-b\int f} \int \frac{ye^{b\int f}}{f\delta(\alpha)}.$$

The integrals occurring in (2.10) do not exist in general. We take the following restriction:

(2.11) $$\frac{y}{f} = \frac{r}{\delta(N)} \in E^*, \quad r \in E, \; N \in Z.$$

Without loss of generality it can be assumed that $\frac{y}{f} \in E$ implies $N=1$.

By introducing

(2.12) $$H_1 = re^{-b\int f}, \quad H_2 = re^{b\int f}$$

we obtain the algebraic integrals

$$\int \frac{H_1}{\delta(\alpha N)}, \quad \int \frac{H_2}{\delta(\alpha N)},$$

existing by the Property 8 of Chapter 1 if and only if either $\alpha N = \frac{N_2}{N_1} N$ is not an integer, or αN is an integer and

$$H_1(\alpha N) = H_2(\alpha N) = 0.$$

Moreover,

$$\int \frac{H_1}{\delta(\alpha N)} = \frac{\left\{-\dfrac{H_1(n)}{\log \dfrac{n}{\alpha N}}\right\}}{\delta(\alpha N)},$$

(2.13)

$$\int \frac{H_2}{\delta(\alpha N)} = \frac{\left\{-\dfrac{H_2(n)}{\log \dfrac{n}{\alpha N}}\right\}}{\delta(\alpha N)},$$

where in the case of $\alpha N \in \mathbf{Z}$ the symbols

$$\frac{H_1(\alpha N)}{\log 1}, \quad \frac{H_2(\alpha N)}{\log 1}$$

denote the number zero.

By substituting (2.13) into (2.10) we have

$$x_p = \frac{1}{2b}\delta\left(\frac{1}{N}\right)\left[e^{-b\int f}\left\{\frac{H_2(n)}{\log \dfrac{n}{\alpha N}}\right\} - e^{b\int f}\left\{\frac{H_1(n)}{\log \dfrac{n}{\alpha N}}\right\}\right]. \tag{2.14}$$

It is easy to see that for $N=1$, i.e. for $\frac{y}{f} \in E$, $x_p \in E$ holds.

In some special cases x_p may be a function also for $N>1$ but we shall not deal with this in this paper.

II. α is irrational.

Then the homogeneous equation has only the trivial solution. However, (2.14) has a meaning since $\delta(\alpha)$ does not occur in it, and an easy substitution shows that it satisfies (2.1). So it holds the following

THEOREM 3. *Let* $f(1)=0$, $b \neq 0$, $y \neq 0$. *Moreover let*

$$\frac{y}{f} = \delta\left(\frac{1}{N}\right) r, \quad r \in E, \ N \in Z,$$

where $\frac{y}{f} \in E$ *implies* $N=1$.

If α *is rational, then* (2.1) *has a particular solution* x_p *in* M *if and only if either* αN *is not an integer, or* αN *is an integer and*

$$H_1(\alpha N) = H_2(\alpha N) = 0,$$

where H_1, H_2 *are given by* (2.11), (2.12).

If α is irrational, then (2.1) *has exactly one solution x_p in M. x_p is given by* (2.14), *and if* $\frac{y}{f}\in E$, *then* $x_p\in E$, *too.*

The general solution of (2.1) *is of the form*

$$x = c_1x_1+c_2x_2+x_p, \quad c_1, c_2\in K$$

provided that the M-operators x_1, x_2, x_p exist.

An interesting consequence of the above results is the following

Alternative Theorem. *Let* $f(1)=0$, $b\neq 0$, $\frac{y}{f}\in E^*$. *If the homogeneous equation* (2.1) *has only the trivial solution in M, then the inhomogeneous equation* (2.1) *has for every $y\in E$ exactly one solution in M. If the homogeneous equation* (2.1) *has a nontrivial solution, then the corresponding inhomogeneous equation has either no solution, or infinitely many solutions in M, according to $y\in E$.*

Every solution of (2.1) *belongs to the ring E^*.*

If $\frac{y}{f}\in E$, *then this Alternative Theorem holds true also in the ring E.*

Remark. Let us observe that if we reject the condition $\frac{y}{f}\in E$, then the Alternative Theorem does not hold in the ring E.

REFERENCES

[1] Fényes, T. and Kosik, P., Az operátortest néhány speciális másodrendű differenciálegyenletéről' II, *Mat. Lapok* **31** (1978/83), 125—134.

[2] Fényes, T. and Szilárd, K., Über diskrete Mikusinskische Operatoren, die auf Grund der Dirichletschen Produktenformel erzeugt werden, *Studia Sci. Math. Hungar.* **11** (1976), 181—199. *MR* **81b**: 44014b.

[3] Fényes, T., On a discrete nonlinear operational differential equation system based on the Dirichlet product, *Studia Sci. Math. Hungar.* **22** (1987), 471—484.

[4] Gesztelyi, E., The application of the operational calculus in the theory of numbers, *Number Theory* (Colloq. Math. Soc. J. Bolyai, Debrecen, 1968), Colloq. Math. Soc. J. Bolyai, Vol. 2, North-Holland, Amsterdam, 1970, 51—104. *MR* **42** # 5922.

[5] Butzer, P. and Schulte, H., *Ein Operatorenkalkül zur Lösung gewöhnlicher und partieller Differenzengleichungssysteme von Funktionen diskreter Veränderlicher und seine Anwendungen,* Forschungsberichte des Landes Nordrhein-Westfalen, Nr. 1557, Westdeutscher Verlag, Köln und Opladen, 1965. *MR* **33** # 511.

[6] Fényes, T., Az operátortest néhány speciális másodrendű differenciálegyenletéről, *Mat. Lapok* **27** (1976/79), 333—354.

(Received July 21, 1986)

MTA MATEMATIKAI KUTATÓ INTÉZETE
P.O. BOX 127
H—1364 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 213—217

A REMARK ON AN m-TH ORDER ALGEBRAIC DIFFERENTIAL EQUATION WITH CONSTANT COEFFICIENTS

TAMÁS FÉNYES

In this remark we shall discuss the algebraic differential equation

$$(1) \qquad D^m(x)+a_{m-1}D^{m-1}(x)+\ldots+a_0x=0, \quad x\in M$$

defined in the discrete Mikusiński operator field M based on the number-theoretical Dirichlet product.

In formula (1) D denotes the symbol of the algebraic derivative, the coefficients of (1) are assumed to be complex numbers. So (1) is a differential equation with constant coefficients in the sense, that for these

$$D(a_k)=0 \qquad k=0,1,\ldots,m-1$$

holds true.

The equation (1) itself is not of much interest, since the solutions of (1) are too special M-operators. However, some higher order algebraic differential equations with non-constant coefficients can be transformed into differential equations with constant coefficients, motivating the discussion of (1). In a forthcoming paper the author will give the operational treatment of some second order equations, which can be transformed to the form (1).

Notations, definitions, and operational rules are given in ([1], [2], [3], [4]): Z, R, K, E, denote the set of the natural numbers, positive rational numbers, the complex numbers, the ring of the real (or complex) valued functions defined in Z, respectively. The ring operations are introduced by the usual addition and the Dirichlet product

$$(2) \qquad ab=\left\{\sum_{\nu/n}a(\nu)b\left(\frac{n}{\nu}\right)\right\}, \quad a,b\in E,\ n=1,2,\ldots,$$

M denotes the field of the Mikusiński operators based on the Dirichlet product.

The operator function $\delta(\alpha)$, $\alpha\in R$ is defined by

$$\delta(\alpha)=\frac{\delta(N_1)}{\delta(N_2)}\in M, \quad \alpha=\frac{N_1}{N_2}, \quad N_1,N_2\in Z,$$

1980 *Mathematics Subject Classification* (1985 Revision). Primary 44A40.
Key words and phrases. Mikusiński's operator calculus, operational differential equations.

where

$$\delta(N) = \{\delta(n, N)\}, \quad N \in Z$$

$$\delta(n, N) = \begin{cases} 0 & \text{for} \quad n \neq N \\ 1 & \text{for} \quad n = N. \end{cases}$$

$$\delta(\alpha)\delta(\beta) = \delta(\alpha\beta) \qquad \alpha, \beta \in R,$$

moreover, $\delta(\alpha) \in E$ if and only if N_2 is a divisor of N_1. $K \subset M$, and the common unit element of K, E, M is $\delta(1) = 1$.

$$D(a) = \{-\log n \cdot a(n)\}, \quad a \in E$$

(3) $$D(x) = \frac{bD(a) - aD(b)}{b^2}, \quad a, b \in E, \quad x = \frac{a}{b} \in M, \, b \neq 0.$$

If $D(x) = 0$ for $x \in M$, then x is an arbitrary complex number. Further

(4) $$D[\delta(\alpha)] = -\log \alpha \, \delta(\alpha) \qquad \alpha \in R.$$

The algebraic integral denoted by $\int$ is the inverse of D.
If $\alpha \in R$, $\alpha \neq 1$, then

(5) $$\int \delta(\alpha) = -\frac{\delta(\alpha)}{\log \alpha} + c, \quad c \in K$$

moreover,

(6) $$\int \delta(1) \text{ does not exist in } M.$$

If the differential equation

$$D(x) - fx = 0 \qquad f \in E$$

has a nontrivial solution x_0 in M, then

$$D(x) - fx = h \qquad h \in E$$

has a solution in M if and only if $\int \frac{h}{x_0}$ exists in M.

In the sequel, the algebraic equation

(7) $$u^m + a_{m-1}u^{m-1} + \ldots + a_1 u + a_0 = 0, \qquad u \in K$$

will be called the characteristic equation of (1).

We prove the following

THEOREM. *Let us consider the algebraic differential equation* (1) *with coefficients* $a_k \in K$ $(a_0 \neq 0)$. *To every real root* u_i *of the characteristic equation* (7) *with*

$$\varrho_i = e^{-u_i} \in R$$

there corresponds the nontrivial solution $\delta(\varrho_i)$ *of* (1). *The general solution in* M *is of the form*

(8) $$x = \sum_i \gamma_i \delta(\varrho_i) \qquad \gamma_i \in K.$$

Let

$$\mu_i = \gamma_i, \quad for \quad \varrho_i \in Z,$$

$$\mu_i = 0, \quad for \quad \varrho_i \notin Z$$

so the general function solution of (1) *is of the form*

$$x = \sum_i \mu_i \delta(\varrho_i). \tag{9}$$

PROOF (by mathematical induction). For $m=1$, the theorem is a special case of a result of [2] (see page 194). Let us assume that the theorem holds for a fixed integer $k>1$ and consider the differential equation

$$D^{k+1}(x)+a_k D^k(x)+\ldots+a_1 D(x)+a_0 x = 0, \tag{10}$$

which can be written as

$$\prod_{j=1}^{k+1} (D-u_j)(x) = \prod_{j=1}^{k} (D-u_j)(D-u_{k+1})(x) = 0 \tag{11}$$

where the numbers u_j are the roots of the corresponding characteristic equation. Let

$$D(x)-u_{k+1}x = y, \tag{12}$$

so we have

$$\prod_{j=1}^{k} (D-u_j)(y) = 0. \tag{13}$$

By our assumption the statement holds for (13). If (13) has only the trivial solution

$$y = 0,$$

then

$$D(x)-u_{k+1}x = 0,$$

and the problem is reduced to the case of $m=1$.

If the general solution of (13) is of the form

$$y = \sum_\tau \gamma_\tau \delta(\varrho_\tau) \qquad \varrho_\tau \in R, \ \gamma_\tau \in K, \tag{14}$$

then we obtain

$$D(x)-u_{k+1}x = \sum_\tau \gamma_\tau \delta(\varrho_\tau). \tag{15}$$

At first let us assume that the homogeneous equation

$$D(z)-u_{k+1}z = 0, \qquad z \in M$$

has a nontrivial solution, then the general solution is given by

$$z = \gamma_{k+1}\delta(\varrho_{k+1}), \quad \varrho_{k+1} = e^{-u_{k+1}} \in R.$$

Consequently, (15) has a solution in M if and only if

$$\int \sum_{\tau} \frac{\gamma_\tau \delta(\varrho_\tau)}{\gamma_{k+1}\delta(\varrho_{k+1})} = \frac{1}{\gamma_{k+1}} \sum_{\tau} \gamma_\tau \int \delta\left(\frac{\varrho_\tau}{\varrho_{k+1}}\right)$$

exists.

Let Q be the set of those indices τ, for which

$$\varrho_\tau \neq \varrho_{k+1} \qquad \tau \in Q.$$

Taking into account (5), (6), we have

$$\sum_{\tau} \gamma_\tau \int \delta\left(\frac{\varrho_\tau}{\varrho_{k+1}}\right) = \sum_{\tau\in Q} \gamma_\tau \int \delta\left(\frac{\varrho_\tau}{\varrho_{k+1}}\right) \tag{16}$$

and then

$$\frac{1}{\gamma_{k+1}} \sum_{\tau\in Q} \gamma_\tau \int \delta\left(\frac{\varrho_\tau}{\varrho_{k+1}}\right) = -\frac{1}{\gamma_{k+1}} \sum_{\tau\in Q} \gamma_\tau \frac{\delta\left(\frac{\varrho_\tau}{\varrho_{k+1}}\right)}{\log \frac{\varrho_\tau}{\varrho_{k+1}}}$$

holds. Applying the method of variation of parameters, we have

$$x = \gamma_{k+1}\delta(\varrho_{k+1}) - \sum_{\tau\in Q} \frac{\gamma_\tau \delta(\varrho_\tau)}{\log \frac{\varrho_\tau}{\varrho_{k+1}}}. \tag{17}$$

Since the numbers γ_τ $(\tau \in Q)$ may be chosen arbitrarily,

$$x = \gamma_{k+1}\delta(\varrho_{k+1}) + \sum_{\tau\in Q} \gamma_\tau \delta(\varrho_\tau)$$

will be obtained. Consequently, (8) holds true for the m-th order equation.

Let us now consider the case where the homogeneous equation

$$D(z) - \mu_{k+1} z = 0$$

has only the trivial zero solution. Then the method of variation of parameters looses its meaning.

However, since (14) is the general solution of (13), it can be easily seen by (11), that it is also the general solution of (10). Consequently, (8) holds true. The last statement of the Theorem is a simple consequence of the two facts that

(i) $\delta(\varrho_i) \in E$ if and only if $\varrho_i \in Z$, and

(ii) $\sum_i \varepsilon_i \delta(\varrho_i) \in E$, for $\varrho_i \notin Z$, $\varepsilon_i \in K$ holds if and only if

$$\varepsilon_i = 0$$

for every i.

REMARK. Let us observe that to a multiple root of the characteristic equation (17) do not correspond two or more linearly independent solutions of (1).

REFERENCES

[1] GESZTELYI, E., The application of the operational calculus in the theory of numbers, *Number Theory* (Colloq. János Bolyai Math. Soc., Debrecen, 1968), Colloq. Math. Soc. J. Bolyai, Vol. 2, North-Holland, Amsterdam, 1970, 51—104. *MR* **42** # 5922.

[2] FÉNYES, T. and SZILÁRD, K., Über diskrete Mikusińskische Operatoren, die auf Grund der Dirichletschen Produktenformel erzeugt werden, *Studia Sci. Math. Hungar.* **11** (1976), 181—199. *MR* **81b**: 44014b.

[3] FÉNYES, T., On a discrete nonlinear operational differential equation system based on the Dirichlet product, *Studia Sci. Math. Hungar.* **22** (1987), 471—484.

[4] FÉNYES, T., On a second order algebraic differential equation, *Studia Sci. Math. Hungar.* **24** (1989), 201—211.

(Received July 21, 1986)

MTA MATEMATIKAI KUTATÓ INTÉZETE
P.O. BOX 127
H—1364 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 219—238

TRANSFORMATIONS CONFORMES DANS LES CORPS HEDERIQUES

YVETTE FENEYROL-PERRIN

Abstract

We show that two open sets of an hederic field — i.e. an algebraically closed, complete valued field whose valuation ring contains a strictly decreasing sequence of prime ideals with the zero ideal for intersection — are conformally equivalent iff the algebras of analytic functions on these open sets are isomorphic. We show, too, that the only conformal mappings between two open sets are Möbius transformations. Where the hederic field and its residual field have zero characteristic, analytic functions satisfy a mean value theorem and Taylor's formulas.

Je remercie Labib Haddad pour les nombreuses conversations qui ont stimulé ce travail.

I. Introduction

Convenons d'appeler corps hédérique tout corps valué, complet, algébriquement clos, dont l'anneau de valuation possède une suite strictement décroissante d'idéaux premiers ayant pour intersection l'idéal nul.

Dans deux articles antérieurs ([5] et [6]) L. Haddad et Y. Perrin ont caractérisé les ouverts D d'un corps hédérique pour lesquels l'algèbre $\mathscr{F}(D)$ des fonctions analytiques sur D est un anneau de Bezout, et ils ont montré que les ouverts d'un tel corps se répartissent en deux classes: Pour les ouverts D de la première classe ou P-ouverts, l'anneau $\mathscr{F}(D)$ est principal, pour les ouverts D de la seconde classe, l'anneau $\mathscr{F}(D)$ n'est pas bezoutien.

Il était donc naturel de se demander dans quelle mesure l'algèbre $\mathscr{F}(D)$ caractérise l'ouvert D. Nous montrons ici que, D et Δ étant deux ouverts d'un corps hédérique K, les algèbres $\mathscr{F}(D)$ et $\mathscr{F}(\Delta)$ sont K-isomorphes si et seulement si D et Δ sont conformément équivalents, c'est-à-dire s'il existe une fonction bianalytique de D sur Δ. Nous montrons d'autre part que les seules fonctions bianalytiques sont les fonctions homographiques. Nous étudions également la dérivation dans les algèbres $\mathscr{F}(D)$ et obtenons entre autre, un théorème des accroissements finis et des formules de Taylor, lorsque K et son corps résiduel sont de caractéristique nulle. Ces résultats attestent que les fonctions analytiques dans un corps hédérique se comportent pour une part comme les fonctions analytiques de la variable complexe et pour une autre part comme les fonctions analytiques de la variable réelle.

1.1. Notations. Soit K un corps hédérique, nous désignons par Γ son groupe des valeurs, par $|\,.\,|$ sa valeur absolue, par A son anneau de valuation, par M son idéal

---

1980 *Mathematics Subject Classification.* Primary 12B40, 30G05, 32D25; Secondary 46P05.
Key words and phrases. Valued field, analytic functions, conformal mapping.

maximal et, pour tout idéal premier P de A, par A_P l'anneau local de A en P et k_P le corps résiduel $A_{P/P}$.

Pour tout $a \in K$ et tout $r \in \Gamma$, nous notons $D(a, r^-)$ (resp. $D(a, r^+)$) le disque non circonférencié (resp. circonférencié) de centre a et de rayon r:

$$D(a, r^-) = \{x \in K\colon |x-a| < r\}$$

$$D(a, r^+) = \{x \in K\colon |x-a| \leqq r\}.$$

1.2. DÉFINITIONS. Soit $\mathscr{P}$ l'ensemble des idéaux premiers non nuls de A. Nous appellerons *suite cofinale dans* $\mathscr{P}$ toute suite $(P_n)_{n \in \mathbf{N}}$ strictement décroissante dont l'intersection est l'idéal nul. Rappelons les définitions de P-ouvert et de P-intérieur, ainsi que celle de P-fraction ([4], p. 200 et [5]).

Soit $P \in \mathscr{P}$. Une partie D de K est un *P-ouvert* lorsque $D \subset A_P$ et $a + P \subset D$ pour tout $a \in D$.

Soit D une partie quelconque de K. Nous appelons *P-intérieur* de D le plus grand P-ouvert contenu dans D et nous le notons $D(P)$. Un élément $f \in K(X)$ est une *P-fraction* si $f(x) \in A_P$ pour tout $x \in P$. On sait que cette condition est équivalente à a suivante: $f(X)$ peut se mettre sous la forme irréductible

$$f(X) = \frac{R(X)}{\prod_i \left(1 - \dfrac{X}{q_i}\right)}$$

où $R(X) \in A_P[X]$ et $q_i \notin P$ quel que soit i.

Pour les définitions d'éléments analytiques et de fonctions analytiques sur les corps hédériques nous renvoyons aux travaux [2] et [4].

Soit D un ouvert quelconque de K, nous notons $H(D)$ l'ensemble des éléments analytiques sur D et $\mathscr{F}(D)$ l'ensemble des fonctions analytiques sur D. Étant donnés une suite cofinale $(P_n)_{n \in \mathbf{N}}$ dans $\mathscr{P}$ et f un élément analytique sur un ouvert D, on appellera *suite approximante* de f toute suite $(f_n)_{n \in \mathbf{N}}$ de fractions rationnelles sans pôle dans D telle que pour n assez grand f_n soit une P_n-fraction, et $f(x) - f_n(x) \in P_n$ pour tout $x \in D$.

On sait que tout élément analytique sur un P-ouvert contenant 0 admet une suite approximante ([5], lemme (4.1) et (4.2)). Rappelons enfin une propriété des fonctions analytiques qui sera essentielle pour toute la suite.

1.3. PROPOSITION ([2], Première partie, chap. IV, § 2, proposition 4, p. 45 et [3]). *Soit D un ouvert quelconque et $f \in \mathscr{F}(D)$. Il existe une suite $(R_n)_{n \in \mathbf{N}}$ de fractions rationnelles sans pôle dans D et une fonction g inversible dans $\mathscr{F}(D)$ telles que:*

i) *le produit $\prod R_n$ converge dans D vers une fonction $\mathbf{Z} \in \mathscr{F}(D)$;*

ii) *$f = \mathbf{Z}g$;*

et

iii) *$f(x) = 0$ si et seulement s'il existe n tel que $R_n(x) = 0$.*

II. Les espaces topologiques $H(D)$ et $\mathscr{F}(D)$. Topologie sur $H(D)$

Pour tout $f \in H(D)$ et tout $\alpha \in \Gamma$, on pose

$$V(f, \alpha) = \{g \in H(D) \colon |f(x) - g(x)| < \alpha \text{ pour tout } x \in D\}.$$

La famille $\{V(f, \alpha)\}_{\alpha \in \Gamma}$ constitue un système fondamental de voisinages de f pour une topologie sur $H(D)$. L'espace $H(D)$ muni de cette topologie est un K-espace vectoriel topologique complet. Si D est un P-ouvert, pour tout $f \in H(D)$, il existe $\alpha \in \Gamma$ tel que $|f(x)| < \alpha$ quel que soit $x \in D$ ([2], première partie, chap. II, proposition 2, p. 31). Il en résulte, dans ce cas, que $H(D)$ est une K-algèbre topologique. Dans tous les cas, la sous-algèbre $K(D)$ des fractions rationnelles sans pôle dans D est dense dans $H(D)$ en raison de la définition même des éléments analytiques.

Topologie sur $\mathscr{F}(D)$.

Rappelons les deux résultats suivants:

2.1. Proposition ([2], p. 29 et [4], p. 63). *Si D est un P-ouvert alors $H(D) = \mathscr{F}(D)$.*

2.2. Proposition ([2], p. 30). *Soit D un ouvert quelconque de K et, pour tout $P \in \mathscr{P}$, soit $D(P)$ son P-intérieur. La famille $\{\mathscr{F}(D(P))\}_{P \in \mathscr{P}}$ est un système projectif d'algèbres et*

$$\mathscr{F}(D) = \varprojlim \mathscr{F}(D(P)).$$

L'ouvert $D(P)$ est un P-ouvert, donc $\mathscr{F}(D(P)) = H(D(P))$. On met sur $H(D(P))$ la topologie précédemment définie et sur $\mathscr{F}(D)$ la topologie limite projective de ces topologies.

2.3. Notations. Soient D un ouvert de K et $f \in \mathscr{F}(D)$. Pour toute partie B de D et tout idéal $P \in \mathscr{P}$, on pose

$$V(f, B, P) = \{g \in \mathscr{F}(D) \colon g(x) - f(x) \in P \quad \text{pour tout} \quad x \in B\}.$$

2.4. Remarque. Soient $P \in \mathscr{P}$ et $Q \in \mathscr{P}$ tels que $P \subset Q$. Alors $D(P) \supset D(Q)$, donc

$$V(f, D(P), P) \subset V(f, D(P), Q) \subset V(f, D(Q), Q).$$

On en déduit immédiatement le résultat suivant.

2.5. Proposition. *Un système fondamental de voisinages de 0 dans $\mathscr{F}(D)$ est constitué par la famille $\{V(0, D(P), P)\}_{P \in \mathscr{P}}$. Soit $(P_n)_{n \in \mathbf{N}}$ une suite cofinale dans $\mathscr{P}$. Alors la famille $\{V(0, D(P_n), P_n)\}_{n \in \mathbf{N}}$ est également un système fondamental de voisinages de 0. L'espace $\mathscr{F}(D)$ est donc métrisable.*

2.6. Proposition. *Soit D un ouvert quelconque de K. La sous-algèbre $K(D)$ des fractions rationnelles sans pôles dans D est dense dans $\mathscr{F}(D)$.*

DÉMONSTRATION. On suppose que $0 \in D$. Soit $f \in \mathscr{F}(D)$. D'après le lemme (4.1) de [6], il existe un idéal premier Q de A tel que f soit une Q-fonction c'est-à-dire

$$f(x) \in A_Q \quad \text{pour tout} \quad x \in Q \quad (\text{et } Q \subset D).$$

D'autre part, on sait ([6], lemme (4.2)) que si une fonction analytique sur un ouvert D est une P-fonction, si $D(P)$ est le P-intérieur de D, il existe une fraction rationnelle h sans pôles dans D telle que

$$f(x) - h(x) \in P \quad \text{pour tout} \quad x \in D(P).$$

D'après la remarque (2.4), tout voisinage de f contient un voisinage de la forme $V(f, D(P), P)$ où $P \subset Q$. Puisque f est une Q-fonction, elle est a fortiori une P-fonction, donc il existe $h \in K(D) \cap V(f, D(P), P)$.

III. Continuité de la dérivation dans $\mathscr{F}(D)$

3.1. LEMME. *Soient* $P \in \mathscr{P}$ *et* $R(X) = \dfrac{S(X)}{Q(X)}$ *une P-fraction où* $S(X) \in A_P[X]$ *et*

$$Q(X) = \prod_i \left(1 - \frac{X}{q_i}\right), \qquad q_i \in K \setminus P, \quad \textit{pour tout } i.$$

Il existe alors un P-polynôme $T(X)$ *tel que* $R(X) - R(0) = XR'(0) + X^2 \dfrac{T(X)}{Q(X)}$.

DÉMONSTRATION. On peut se ramener au cas où $R(0) = 0$.

1° — Cas d'un P-polynôme.

Si $R(X) = a_1 X + \ldots + a_n X^n$ est un P-polynôme, alors

$$R(X) - XR'(0) = a_2 X^2 + \ldots + a_n X^n = X^2 T(X) \quad \text{et} \quad T(X) \in A_P[X].$$

2° — Cas général: on a

$$R'(X) = \frac{S'(X)Q(X) - S(X)Q'(X)}{Q^2(X)} \quad \text{et} \quad Q(0) = 1$$

$$R'(0) = S'(0)$$

$$R(X) - XR'(0) = \frac{S(X) - S'(0)XQ(X)}{Q(X)}$$

$$= \frac{S(X) - S'(0)X - S'(0)X(Q(X) - 1)}{Q(X)}$$

$$S(X) - S'(0)X = X^2 T_1(X) \quad \text{où} \quad T_1(X) \in A_P[X]$$

$$Q(X) - 1 = Q(X) - Q(0) = XT_2(X) \quad \text{où} \quad T_2 \in A_P[X]$$

car $Q(X) \in A_P[X]$.

Donc

$$R(X)-XR'(0)=\frac{X^2[T_1(X)-S'(0)\,T_2(X)]}{Q(X)}=X^2\frac{T(X)}{Q(X)}$$

et $T\in A_P[X]$ puisque $S'(0)\in A_P$.

3.2. Corollaire. *Soit $R(X)$ une P-fraction, alors*

$$\frac{R(x)-R(0)}{x}-R'(0)\in P \quad \textit{pour tout} \quad x\in P\setminus\{0\}.$$

Démonstration. D'après le lemme (3.1),

$$\frac{R(x)-R(0)}{x}-R'(0)=x\frac{T(x)}{Q(x)}$$

et $\dfrac{T(X)}{Q(X)}$ est une P-fraction. Donc $x\dfrac{T(x)}{Q(x)}\in P$ si $x\in P\setminus\{0\}$. Pour toute la fin de ce paragraphe, on se fixe une suite cofinale $(P_n)_{n\in\mathbf{N}}$ dans $\mathscr{P}$.

3.3. Lemme. *Soient D un P-ouvert, $f\in\mathscr{F}(D)$, $(f_n)_{n\in\mathbf{N}}$ une suite approximante de f sur D. On suppose que f est une Q-fonction. Alors*

i) *f est dérivable sur D et sa dérivée $f'\in\mathscr{F}(D)$,*

ii) *$(f_n')_{n\in\mathbf{N}}$ est une suite approximante de f',*

et

iii) *f' est une Q-fonction.*

Démonstration. On montrera successivement que
a) la suite $(f_n')_{n\in\mathbf{N}}$ converge uniformément sur D vers une Q-fonction g;
b) $g(0)=f'(0)$;
c) $g(a)=f'(a)$ pour tout $a\in D$.

a) Il existe $N\in\mathbf{N}$ tel que pour tout $n>N$, on ait:

$$P_n\subset Q, \quad \text{et} \quad P_n\subsetneqq P$$

$$f_n \text{ est une } P_n\text{-fraction et}$$

$$f_n(x)-f_{n+1}(x)\in P_n, \quad \text{pour tout} \quad x\in D.$$

Soit k_{P_n} le corps résiduel A_{P_n/P_n}. Puisque $P_n\subsetneqq P$, on en déduit que les éléments $\bar{f}_n$ et $\bar{f}_{n+1}$ de $k_{P_n}(X)$ sont égaux, mais alors les fractions dérivées $(\bar{f}_n)'$, et $(\bar{f}_{n+1})'$ éléments de $k_{P_n}(X)$, le sont également. Les fractions $f_n'(X)$ et $f_{n+1}'(X)$ sont des P_n-fractions. On a donc, dans $k_{P_n}(X)$

$$\overline{f_n'}=(\bar{f}_n)'=(\bar{f}_{n+1})'=\overline{f_{n+1}'}.$$

Il en résulte que, pour tout $x\in D\subset D(P_n)$,

$$f_n'(x)-f_{n+1}'(x)\in P_n.$$

La suite $(f_n'(x))$ est une suite de Cauchy. Notons $g(x)$ sa limite. On a donc

$f_n'(x)-g(x)\in P_n$ pour tout $n>N$ et tout $x\in D$. Or f_n est une Q-fraction donc f_n' est également une Q-fraction, donc g est une Q-fonction.

b) On peut supposer que $f(0)=0$. Soit $R\in\mathscr{P}$, $R\subset P\cap Q$ et soit $x\in R$, $x\neq 0$. On a

$$\left[\frac{f(x)}{x}-g(0)\right]=\left[\frac{f(x)}{x}-\frac{f_n(x)-f_n(0)}{x}\right]+\left[\frac{f_n(x)-f_n(0)}{x}-f_n'(0)\right]+[f_n'(0)-g(0)].$$

Soit $n_0>N$ tel que $P_{n_0}\subset R$. Il existe $n>n_0$ tel que $x\notin P_n$, alors pour un tel n on a:

$$\frac{f(x)}{x}-\frac{f_n(x)-f_n(0)}{x}\in P_n\subset R,$$

$$f_n'(0)-g(0)\in P_n$$

et

$$\frac{f_n(x)-f_n(0)}{x}-f_n'(0)\in R.$$

En effet f_n est une R-fraction puisque pour tout $n>N$, f_n est une Q-fraction et que l'on a $R\subset Q$. On peut donc lui appliquer le corollaire (3.2). Finalement, on a, pour tout $x\in R$ $\dfrac{f(x)}{x}-f(0)\in R$ donc $g(0)=f'(0)$.

c) Soit $a\in D$, on considère la fonction $F(x)=f(x+a)$ et on lui applique le résultat précédent, mais avec un idéal Q qui n'est plus nécessairement le même.

3.4. Corollaire. *Soient D un P-ouvert, $f\in\mathscr{F}(D)$, $(f_n)_{n\in\mathbf{N}}$ une suite approximante de f sur D. Alors*

i) *f est indéfiniment dérivable sur D;*

ii) *la $k^{\text{ème}}$ dérivée $f^{(k)}$ de f est une fonction analytique sur D, admettant la suite $(f_n^{(k)})_{n\in\mathbf{N}}$ comme suite approximante et*

iii) *si f est une Q-fonction, alors $f^{(k)}$ est aussi une Q-fonction.*

3.5. Corollaire. *Soit $(f_n)_{n\in\mathbf{N}}$ une suite de fonctions analytiques sur le P-ouvert D qui converge uniformément sur D vers f. La suite des dérivées $(f_n')_{n\in\mathbf{N}}$ converge alors uniformément sur D vers f'.*

Démonstration. Pour n assez grand, il existe une fraction rationnelle $g_n(X)$ sans pôle dans D telle que

$$f_n(x)-g_n(x)\in P_n.$$

D'après le lemme (3.3), $f_n'(x)-g_n'(x)\in P_n$ pour tout $x\in D$. La suite $(g_n)_{n\in\mathbf{N}}$ converge uniformément sur D vers f. Donc la suite $(g_n')_{n\in\mathbf{N}}$ converge uniformément sur D vers f' (d'après le lemme (3.3)) et par conséquent la suite $(f_n')_{n\in\mathbf{N}}$ converge uniformément sur D vers f'.

On en déduit immédiatement le corollaire suivant:

3.6. Corollaire. *Soient D un ouvert quelconque de K, et $(f_n)_{n\in\mathbf{N}}$ une suite de fonctions analytiques sur D, qui converge dans $\mathscr{F}(D)$ vers une fonction f. La suite des dérivées $(f_n')_{n\in\mathbf{N}}$ converge alors dans $\mathscr{F}(D)$ vers f'.*

IV. Théorème des accroissements finis et formules de Taylor

Les résultats obtenus dans le paragraphe précédent ont leur équivalent en analyse complexe. On va maintenant donner des résultats qui rapprochent au contraire les fonctions analytiques sur un corps hédérique, des fonctions analytiques réelles: théorème des accroissements finis, formules de Taylor. Soulignons que le théorème des accroissements finis est également vrai pour les corps valués de rang un, autrement dit en analyse ultramétrique classique (voir ci-dessous les remarques (4.6)).

Rappelons tout d'abord quelques propriétés bien connues de la fonction de valuation d'un polynôme à coefficients dans un corps valué algébriquement clos donc hensélien ([7] et [2], p. 19).

4.2. Rappel. Soit $P(X)\in K[X]$. $P(X)=a_0+a_1X+\ldots+a_nX^n$. La fonction de valuation de $P(X)$ est définie sur $\Gamma\cup\{0\}$ par

$$M_P(r) = \sup_i |a_i|\, r^i.$$

Notons

$$N_P(r) = \sup\{i\colon\ |a_i|r^i\} = M_P(r)\}$$

$$n_P(r) = \inf\{i\colon\ |a_i|r^i\} = M_P(r)\}.$$

Alors sur la circonférence $C(r)=\{x\in K\colon\ |x|=r\}$,

1) $P(X)$ admet $N_P(r)-n_P(r)$ zéros, chaque zéro étant compté un nombre de fois égal à son ordre de multiplicité.

2) $|P(x)|\leqq M_P(r)$ pour tout $x\in C(r)$.

3) $N_P(r)=n_P(r)$ si et seulement si $|P(x)|=M_P(r)$, pour tout $x\in C(r)$ et

4) si $N_P(r)>n_P(r)$ et si $y\in C(r)$ est tel que $|P(y)|<M_P(r)$, alors il existe $x\in C(r)$ tel que $P(x)=0$ et $|x-y|<r$.

4.2. Définition. Un polynôme $P(X)\in A[X]$ est dit *primitif* s'il admet au moins un coefficient de valeur absolue égale à un, c'est-à-dire encore si le polynôme correspondant $\bar{P}(X)$ dans $k(X)$ est non nul.

On déduit immédiatement de (4.1) le lemme suivant:

4.3. Lemme. *Pour qu'un polynôme primitif admette au moins un zéro de valeur absolue égale à un il faut et il suffit qu'il ait au moins deux coefficients de valeur absolue égale à un.*

4.4. Lemme [on suppose K et k de caractéristique nulle]. *Soit $R(X)$ une fraction rationnelle sans pôle dans A, telle que:*

$$R(0) = R(1) = 0.$$

Alors il existe au moins un $x\in K$ tel que

$$R'(x) = 0 \quad et \quad |x| = 1.$$

Démonstration. Soit $R(X)=\dfrac{P(X)}{Q(X)}$ irréductible où $Q(0)=1$. On peut supposer $R\neq 0$ et en divisant par un coefficient de $P(X)$ de plus grande valeur absolue se ramener au cas où $P(X)$ est un polynôme primitif. On pose

$$P(X) = a_0 + a_1 X + \ldots + a_n X^n.$$

Puisque $P(0) = P(1) = 0$, il vient $a_0 = 0$ et $a_1 + \ldots + a_n = 0$ donc $P(X)$ possède au moins deux coefficients de valeur absolue égale à un. Puisque $Q(X)$ ne s'annule pas sur A, il s'écrit

$$Q(X) = 1 + b_1 X + \ldots + b_q X^q \quad \text{où} \quad b_j \in M \quad \text{pour tout} \quad j \in \{1, \ldots, q\}.$$

D'autre part,

$$R'(X) = \frac{P'(X)Q(X) - P(X)Q'(X)}{Q^2(X)}.$$

Dans $k[X]$ on a

$$\overline{P'}(X)\overline{Q}(X) - \overline{P}(X)\overline{Q'}(X) = \overline{P'}(X).$$

Or $\overline{P}(X)$ a au moins deux coefficients non nuls, d'indices différents de zéro, k est de caractéristique nulle, donc $\overline{P'}(X)$ a au moins deux coefficients non nuls. D'après le lemme précédent le polynôme primitif $P'(X)Q(X) - P(X)Q'(X)$ a au moins un zéro x de valeur absolue égale à un.

4.5. Théorème des accroissements finis pour les fractions [On suppose K et k de caractéristique nulle]. *Soient a et b deux points distincts de K, $R(X)$ une fraction rationnelle. On suppose que $R(X)$ n'a pas de pôle dans le disque $D(a, |b-a|^+)$. Il existe alors au moins un $x \in K$ tel que*

$$R(b) - R(a) = (b-a) R'(a+x) \quad \text{et} \quad |x| = |b-a|.$$

Démonstration. On précède classiquement: on applique le lemme (4.4) à

$$S(X) = R\big(a + (b-a)X\big) - R(a) - [R(b) - R(a)]\, X.$$

4.6. Remarques.

4.6.1. Le théorème (4.5) est vrai dans tout corps algébriquement clos muni d'une valuation de rang quelconque. Il se généralise facilement aux éléments analytiques lorsque le corps K est muni d'une valuation de rang un. On a alors le résultat suivant:

4.6.2. Théorème des accroissements finis pour les corps valués de rang un. *On se donne un corps valué L de rang un, complet et algébriquement clos. On suppose que L est de caractéristique nulle ainsi que son corps résiduel. Soient D un ouvert de L et $f \in H(D)$. Soient a et b deux points distincts dans D tels que $D(a, |b-a|^+) \subset D$. Il existe alors au moins un $x \in D$ tel que*

$$f(b) - f(a) = (b-a) f'(a+x) \quad \text{et} \quad |x| = |b-a|.$$

Démonstration. On se ramène par une homographie au cas où $a = 0$, $b = 1$ et l'on suppose que $f(0) = f(1) = 0$. Dans ces conditions $f \in H(D(0, 1^+))$. On sait que f est la somme d'une série de Taylor $\sum_{n=0}^{+\infty} a_n X^n$ convergente sur $D(0, 1^+)$, et que la fonction de valuation M_f d'une telle série satisfait à la proposition (4.1) ([8], § 2,

p. 102). Puisque $f(0)=f(1)=0$, on a $a_0=0$ et $\sum_{i=0}^{+\infty} a_i=0$, donc il existe au moins deux indices non nuls i et j distincts tels que $|a_i|=|a_j|=M_f(1)$. Pour tout $x\in D(0, 1^+)$ on a

$$f'(x) = \sum_{n=1}^{+\infty} na_n x^{n-1}$$

donc $M_{f'}(1)=\sup_n |n||a_n|=\sup_n |a_n|$ puisque le corps résiduel est de caractéristique nulle; il existe donc deux indices i et j tels que $M_{f'}(1)=|ia_i|=|ja_j|$ ce qui assure l'existence d'un zéro pour f' sur la circonférence $|x|=1$.

Nous allons généraliser le théorème (4.5) aux fonctions analytiques et pour cela nous utiliserons les deux lemmes suivants:

4.7. LEMME. *Soient D un P-ouvert et $f\in\mathscr{F}(D)$. On suppose que f ne s'annule pas sur D. Alors,*

(i) *Il existe $\alpha\in\Gamma$ tel que $|f(x)|>\alpha$ pour tout $x\in D$.*

(ii) *Dans tout disque $D(a, r^+)$ (resp. $D(a, r^-)$) contenu dans D, $|f|$ est constante.*

DÉMONSTRATION. (i) Puisque f ne s'annule pas sur D, $1/f$ est une fonction analytique sur D ([2], prop. 6, chap. II, p. 32) et comme D est un P-ouvert, la fonction $1/f$ est élément analytique (proposition (2.1)). Or tout élément analytique sur un P-ouvert est borné donc il existe $\beta\in\Gamma$ tel que $|1/f(x)|<\beta$, pour tout $x\in D$.

(ii) Démontrons le résultat pour un disque $D(a, r^+)$; pour un disque non circonférencié $D(a, r^-)$ le résultat s'en déduira immédiatement puisque $D(a, r^-)= =\bigcup_{\varrho<r} D(a, \varrho^+)$.

Par une transformation homographique on se ramène au cas où $D(a, r^+)=A$ et on suppose que $f(0)=1$. Soit $(P_n)_{n\in\mathbf{N}}$ une suite cofinale dans $\mathscr{P}$ et $(f_n)_{n\in\mathbf{N}}$ une suite approximante de f sur A. Pour n assez grand on a d'une part $f(x)-f_n(x)\in\mathscr{P}_n$ quel que soit $x\in A$ et d'autre part $f(x)\notin\mathscr{P}_n$ puisque $|f(x)|>\alpha$, donc $f_n(x)$ ne s'annule pas dans A.

Posons $f_n(x)=\dfrac{P_n(X)}{Q_n(X)}$ avec $Q_n(0)=1$. Puisque $P_n(X)$ et $Q_n(X)$ ne s'annulent pas dans A, on a d'après (4.1), pour tout $x\in A$,

$$|P_n(x)| = |P_n(0)| \quad \text{et} \quad |Q_n(x)| = |Q_n(0)| = 1,$$

donc $|f_n(x)|=|f_n(0)|$. Or

$$|f(x)| = \lim_{n\to\infty} |f_n(x)| = \lim_{n\to\infty} |f_n(0)| = |f(0)|$$

pour tout $x\in A$.

4.8. LEMME. *Soient D un P-ouvert, $(f_n)_{n\in\mathbf{N}}$ une suite de fonctions analytiques sur D qui converge uniformément vers f sur D, $(x_n)_{n\in\mathbf{N}}$ une suite d'éléments de D. On suppose que $f_n(x_n)_{n\in\mathbf{N}}$ converge vers u. Il existe alors un point $x\in D$ tel que $f(x)=u$.*

DÉMONSTRATION. Supposons que pour tout $x \in D$ on ait $f(x)-u \neq 0$. Il existe alors $\alpha \in \Gamma$ tel que $|f(x)-u| > \alpha$ pour tout $x \in D$ (lemme (4.7)). Or

$$\begin{aligned} |f(x_n)-u| &= |f(x_n)-f_n(x_n)+f_n(x_n)-u| \\ &\leq \operatorname{Max}(|f(x_n)-f_n(x_n)|, |f_n(x_n)-u|). \end{aligned}$$

Par hypothèse $\lim_{n\to\infty}(f_n(x_n)-u)=0$, et d'autre part $\lim_{n\to\infty}(f(x_n)-f_n(x_n))=0$ puisque la suite (f_n) converge uniformément vers f sur D, donc $\lim_{n\to\infty}(f(x_n)-u)=0$, ce qui contredit $|f(x_n)-u| > \alpha$.

4.9. THÉORÈME des accroissements finis [On suppose K et k de caractéristique nulle]. *Soient D un ouvert quelconque de K, f une fonction analytique sur D, a et b deux éléments distincts de D. On suppose que $D(a, |b-a|^+) \subset D$. Il existe alors au moins un x tel que*

$$f(b)-f(a) = (b-a)f'(a+x) \quad et \quad |x| = |b-a|.$$

DÉMONSTRATION. Soit $P \in \mathscr{P}$ tel que $P \subset D(a, |b-a|^+)$. Le disque $U = D(a, |b-a|^+)$ est un P-ouvert. Ainsi f est un élément analytique sur U et il existe une suite (f_n) de fractions rationnelles sur K sans pôle dans U qui converge uniformément vers f sur U. D'après le corollaire (3.5) la suite des dérivées (f_n') converge uniformément sur U vers la dérivée f'.

D'après le théorème (4.5), pour chaque n, il existe x_n tel que

$$f(a+x_n) = \frac{f_n(b)-f_n(a)}{b-a} = u_n \quad \text{et} \quad |x_n| = |b-a|.$$

Or la suite $(u_n)_{n\in\mathbf{N}}$ converge vers $\dfrac{f(b)-f(a)}{b-a}$, et la circonférence $C(a, |b-a|) = \{y\colon |y-a| = |b-a|\}$ est un P-ouvert. D'après le lemme (4.8), il existe donc $x \in C(a, |b-a|)$ tel que

$$f'(a+x) = \frac{f(b)-f(a)}{b-a}.$$

4.10. COROLLAIRE [on suppose K et k de caractéristique nulle]. *Une fonction analytique sur un ouvert D quelconque est constante si et seulement si sa dérivée est nulle sur D.*

DÉMONSTRATION. Soit $f \in \mathscr{F}(D)$ telle que $f'(x)=0$ pour tout $x \in D$, et soit $D(a, r^+)$ un disque contenu dans D. D'après le théorème des accroissements finis, f est constante sur $D(a, r^+)$. Mais f vérifie «le principe du prolongement analytique» ([2], Théorème 2, p. 22 et [4], Théorème (4.2), p. 62). Donc f est constante sur D.

4.11 FORMULE de Taylor—Lagrange [on suppose K et k de caractéristique nulle]. *Soient D un ouvert quelconque de K, $f \in \mathscr{F}(D)$, a et b deux éléments distincts de D. On suppose que le disque $D(a, |b-a|^+)$ est contenu dans D. Alors, pour tout $n \in \mathbf{N}$ il existe $x \in D$ tel que*

$$f(b) = \sum_{k=1}^{n} \frac{(b-a)^k}{k!} f^{(k)}(a) + \frac{(b-a)^{n+1}}{(n+1)!} f^{(n+1)}(a+x) \quad et \quad |x| = |b-a|.$$

DÉMONSTRATION. On procède classiquement: on pose

$$\varphi(t) = f(t) + \sum_{k=1}^{n} \frac{(b-t)^k}{k!} f^{(k)}(t)$$

et

$$\psi(t) = \varphi(t) + B \frac{(b-t)^{n+1}}{(n+1)!},$$

B étant tel que $\psi(b) = \psi(a)$. La fonction ψ est analytique sur le disque $D(a, |b-a|^+)$ et

$$\psi'(t) = \frac{(b-t)}{n!} [f^{(n+1)}(t) - B].$$

D'après le théorème des accroissements finis, il existe x tel que

$$\psi'(a+x) = 0 \quad \text{et} \quad |x| = |b-a|$$

c'est-à-dire tel que $B = f^{(n+1)}(a+x)$ or

$$B \frac{(b-a)^{n+1}}{(n+1)!} = f(b) - \sum_{k=0}^{n} \frac{(b-a)^k}{k!} f^{(k)}(a).$$

4.12. FORMULE de Taylor—Young [On suppose K et k de caractéristique nulle].

Soient $P \in \mathscr{P}$ et f une fonction analytique sur P. Il existe alors un idéal $Q \in \mathscr{P}$ contenu dans P tel que f soit une Q-fonction, et il existe une Q-fonction g analytique sur P telle que

$$f(x) = \sum_{k=0}^{n} \frac{x^k}{k!} f^{(k)}(0) + \frac{x^{(n+1)}}{(n+1)!} g(x), \quad \textit{pour tout} \quad x \in P.$$

DÉMONSTRATION. La fonction h définie sur P par $h(x) = f(x) - \sum_{k=0}^{n} \frac{x^k}{k!} f^{(k)}(0)$ est analytique sur P et s'annule en 0 ainsi que ses n premières dérivées. Il existe donc une fonction g analytique sur P telle que $h(x) = \frac{x^{n+1}}{(n+1)!} g(x)$. On a ainsi $g(0) = f^{(n+1)}(0)$. De plus, d'après la formule de Taylor—Lagrange (4.11), pour chaque $x \in P \setminus \{0\}$, il existe $\theta \in K$ tel que $h(x) = \frac{x^{n+1}}{(n+1)!} f^{(n+1)}(\theta x)$ et $|\theta| = 1$; de sorte que $g(x) = f^{(n+1)}(\theta x)$. Or d'après le corollaire (3.4), $f^{(n+1)}$ est une Q-fonction.

V. Image d'un ouvert par une fonction analytique

5.1. LEMME. *Soit $f(X) = X^p g(X)$ où p est un entier non nul, $g(X)$ une fraction rationnelle non nulle sans pôle dans A, telle que $g(0) = 1$.*

i) *On a alors $f(A) \supset A$.*

ii) *Si g ne s'annule pas sur A, alors $f(A) = A$ et l'équation $f(x) - \alpha = 0$ possède dans A exactement p racines distinctes ou confondues, pour tout $\alpha \in A$.*

DÉMONSTRATION. i) Soit $g(X)=\dfrac{P(X)}{Q(X)}$. Nous pouvons supposer que $P(0)=$ $=Q(0)=1$ et puisque $Q(X)$ ne s'annule pas sur A, d'après (4.1) $Q(X)$ s'écrit

$$Q(X) = 1+b_1X+\ldots+b_qX^q \quad \text{où} \quad b_i\in M$$

$$P(X) = 1+a_1X+\ldots+a_nX^n.$$

Soit $\alpha\in A$. D'après la proposition (4.1), le polynôme

$$R(X)=X^pP(X)-\alpha Q(X)=-\alpha-\alpha b_1X-\ldots-\alpha b_{p-1}X^{p-1}+X^p+\ldots+(a_n-\alpha b_{n+p})X^{n+p}$$

a au moins un zéro dans le disque $D(0, r^+)$ où $r^p=|\alpha|$.

ii) On suppose maintenant que g ne s'annule pas sur A. Alors $|g(x)|=1$ pour tout $x\in A$ (Lemme (4.7)) donc $f(A)\subset A$, donc $f(A)=A$.

Puisque $P(X)$ ne s'annule pas sur A, les coefficients a_i sont dans M. Soient $\alpha\in A$ et $r\in\Gamma$. On a, en utilisant les notations de la proposition (4.1)

$$\begin{array}{lll} \text{si} & r^p<|\alpha| & N_R(r)=n_R(r)=0, \\ \text{si} & r^p=|\alpha| & N_R(r)=p;\ n_R(r)=0, \\ \text{si} & 1\geqq r^p>|\alpha| & N_R(r)=n_R(r)=p, \end{array}$$

et l'on en déduit que $R(X)$ a p zéros dans A. Ces zéros ont pour valeur absolue $|\alpha|^{1/p}$.

5.2. LEMME. *Soit $f(X)=X^pg(X)$ où $p\geqq 1$ et g est une fonction analytique sur A telle que $g(0)=1$.*

i) *on a alors $f(A)\supset A$.*

ii) *si g ne s'annule pas sur A, alors $f(A)=A$.*

iii) *si g ne s'annule pas sur A et si* car $K\neq p$ *alors, pour tout $\alpha\in A\setminus\{0\}$ tel que $|\alpha|<|p|^{1/p}$, l'équation $f(x)-\alpha=0$ possède exactement p racines dans A et qui toutes sont simples.*

DÉMONSTRATION. On considère une suite cofinale $(P_n)_{n\in\mathbf{N}}$ dans $\mathscr{P}$.

i) La fonction g admet dans A une suite approximante $(g_n)_{n\in\mathbf{N}}$ telle que $g_n(0)=1$. La suite $f_n(X)=X^pg_n(X)$ est une suite approximante de f. Soit $\alpha\in A$. D'après le lemme (5.1), il existe une suite $(x_n)_{n\in\mathbf{N}}$ de points de A telle que $f_n(x_n)=\alpha$. D'après le lemme (4.8) il existe $x\in A$ tel que $\alpha=f(x)$.

ii) Si g ne s'annule pas sur A, alors $|g(x)|=1$ pour tout $x\in A$ (lemme (4.7)) donc $f(A)=A$.

iii) D'autre part, g admet une suite approximante $(g_n)_{n\in\mathbf{N}}$ telle que pour n assez grand

$$g_n(0) = 1 \quad \text{et}$$

$$g_n(X)=\frac{P_n(X)}{Q_n(X)} \quad \text{oû} \quad |P_n(x)|=1 \quad \text{et} \quad |Q_n(x)|=1, \quad \text{pour tout} \quad x\in A.$$

Donc $P_n(X)$ et $Q_n(X)$ appartiennent à $1+XM(X)$. D'après le lemme (5.1) la fraction rationnelle

$$h_n(X)=X^p g_n(X)-\alpha$$

où $\alpha\in A$, a exactement p zéros dans A. Chaque zéro est compté un nombre de fois égal à son ordre de multiplicité et a pour valeur absolue $|\alpha|^{1/p}$. Notons $a_{i,n}$ ces zéros et écrivons $h_n(X)$ sous la forme

$$h_n(X) = \prod_{i=1}^{p}\left(1-\frac{X}{a_{i,n}}\right)l_n(X) \quad \text{où} \quad l_n(X)\in M(X).$$

La fraction $l_n(X)$ ne s'annule pas sur A donc $|l_n(x)|=|l_n(0)|=|\alpha|$ pour tout $x\in A$ (lemme 4.7). D'autre part, il existe $n_0\notin \mathbf{N}$ tel que, pour tout $n>n_0$ et tout $x\in A$

$$h_n(x)-h_{n+1}(x)\in P_n.$$

Choisissons n_0 tel que $\alpha\in P_{n_0}$, alors pour $n>n_0$, $a_{i,n}^{-1}\in A_{P_n}$ et l'on a dans $k_{P_n}(X)$ $\bar{h}_n(X)=\bar{h}_{n+1}(X)$. On peut renuméroter les racines $a_{i,n+1}$ de sorte que, pour tout $i=1, \ldots, p$ on ait $\bar{a}_{i,n}=\bar{a}_{i,n+1}$ dans k_{P_n}. De plus on a $\bar{l}_n(X)=\bar{l}_{n+1}(X)$ dans $k_{P_n}(X)$, alors les suites $(a_{i,n})_{n\in\mathbf{N}}$ sont des suites de Cauchy. Notons

$$a_i = \lim_{n\to\infty} a_{i,n}.$$

La suite $(l_n(X))_{n\in\mathbf{N}}$ converge dans A vers un élément analytique l. Ainsi on a pour tout $x\in A$

$$f(x)-\alpha = \prod_{i=1}^{p}\left(1-\frac{x}{a_i}\right)l(x) \quad \text{et} \quad |l(x)| = |\alpha|.$$

Supposons que la caractéristique de K soit différente de p, et que $|\alpha|<|p|^p$. Montrons que pour tout $i=1, \ldots, p$, $f'(a_i)\neq 0$.

$$\begin{aligned} f'(a_i) &= pa_i^{p-1}g(a_i)+a_i^p g'(a_i) \\ &= a_i^{p-1}[pg(a_i)+a_i g'(a_i)]. \end{aligned}$$

Pour tout $x\in A$, $g'(x)=\lim_{n\to\infty} g_n'(x)$ d'après le théorème 3.5. Or pour n assez grand, $|g_n'(x)|\leqq 1$, donc $|g'(a_i)|\leqq 1$, et

$$|pg(a_i)+a_i g'(a_i)| = \operatorname{Max}(|p|,|\alpha|^{1/p}|g'(a_i)|) = |p|,$$

donc

$$|f'(a_i)| = |p|\,|\alpha|^{(p-1/p)} > |\alpha|.$$

5.3. Théorème. *Toute fonction analytique sur un ouvert quelconque de K est ouverte ou constante. Plus précisément: Soient D un ouvert de K, $f\in\mathscr{F}(D)$, $a\in D$ et*

$$f(x)-f(a) = (x-a)^p g(x) \quad \textit{où} \quad g\in\mathscr{F}(D) \quad \textit{et} \quad g(a)\neq 0.$$

Pour chaque disque $D(a, r^+)\subset D$, on a alors

$$f(D(a,r^+))\supset D(f(a),s^+) \quad \textit{où} \quad s = r^p|g(a)|.$$

En particulier, si K est de caractéristique nulle ou $\operatorname{car}(K)>p$ *alors* $g(a)=\dfrac{f^{(p)}(a)}{p!}$ *bien entendu.*

Démonstration. Soient $a \in D$, $D(a, r^+)$ un disque contenu dans D et $\alpha \in K$ tel que $|\alpha| = r$. La fonction $F(x) = f(a+\alpha x) - f(a)$ est analytique sur A et s'annule en 0. Supposons f non constante sur D, alors F s'écrit:

$$F(x) = x^p h(x) \quad \text{où} \quad p \geqq 1,\, h \in H(A) \quad \text{et} \quad h(0) \neq 0 \qquad (\text{prop.}\,(1.3)).$$

Si K est de caractéristique nulle $h(0) = \dfrac{\alpha^p}{p!} f^{(p)}(a)$. D'après le lemme (5.2) on a $F(A) \supset |h(0)|A$, donc $f(D(a, r^+)) \supset D(f(a), s^+)$ où $s = |h(0)| = r^p |g(a)|$.

5.4. Corollaire. *Soit $f \in \mathscr{F}(M)$ telle que $f(X) = X^p g(X)$ où $p \geqq 1$, $g \in \mathscr{F}(M)$ et $g(0) = 1$, alors $f(M) \supset M$. Si g ne s'annule pas sur M, alors $f(M) = M$.*

Démonstration. $M = \bigcup\limits_{r<1} D(0, r^+)$. Le résultat découle alors immédiatement du théorème précédent, et du lemme (4.7).

5.5. Théorème. *Soit f une fonction analytique sur un ouvert quelconque D de K, et soit $a \in D$. Si a est une racine d'ordre p de l'équation $f(x) - f(a) = 0$ et si K est de caractéristique nulle ou* car $K > p$ *il existe un voisinage de a dans lequel f est une fonction p-valente.*

Démonstration. Soit $D(a, r^+)$ un disque contenu dans D. La fonction F définie $F(u) = \dfrac{1}{\gamma}[f(a+ru) - f(a)]$ où $\gamma = r^p \dfrac{f^{(p)}(a)}{p!}$, vérifie les hypothèses du lemme 5.2. Donc si

$$|\alpha| < |p|^p \frac{|f^{(p)}(a)|}{|p!|} r^p$$

l'équation $f(x) - f(a) = \alpha$ a p racines distinctes sur la circonférence de centre a et de rayon $r \left|\dfrac{\alpha}{\gamma}\right|^{1/p}$.

Image d'un ouvert par une fonction analytique injective

5.6. Définition. Nous dirons qu'une fonction analytique f définie sur l'ouvert D de K est *strictement injective* sur D si, pour tout $y \in K$, l'équation $f(x) = y$ a au plus une racine simple dans D.

5.7. Proposition. *Pour qu'une fonction analytique f sur un ouvert D soit strictement injective il faut et il suffit que f soit injective et que sa dérivée ne s'annule pas sur D.*

Démonstration. Soit $a \in D$. La fonction analytique h définie sur D par $h(x) = f(x) - f(a)$ admet a comme zéro donc (prop. (1.3)) il existe $g \in \mathscr{F}(D)$ telle que $h(x) = (x-a)g(x)$ et a est zéro simple de h si et seulement si $g(a) = f'(a)$ est non nul.

5.8. Définition. Soient D et Δ deux ouverts de K et f une application de D dans Δ. On dit que h est *bianalytique* lorsqu'elle est bijective analytique et que son application réciproque est analytique.

5.9. Proposition. *Soient D et Δ deux ouverts de K et f: $D \to K$ une fonction bi-analytique. Alors f est strictement injective.*

Démonstration. Soient $a \in D$ et n l'ordre de multiplicité du zéro a de la fonction h définie par $h(x)=f(x)-f(a)$. On se ramène au cas où $a=0$ et $f(a)=0$. Ainsi $h(x)=x^n g(x)$ où $g \in \mathscr{F}(D)$ et $g(0)\neq 0$. Posons $l=h^{-1}$ et $y=h(x)$; Ainsi $y==x^n g(x)=(l(y))^n g\circ l(y)$ et, puisque $l(0)=0$, il existe $k \in \mathscr{F}(D)$ tel que

$$l(y) = yk(y), \quad \text{donc} \quad y = y^n(k(y))^n g\circ l(y)$$

donc $k^n(y) g\circ l(y)=y^{1-n}$. Or, la fonction $k^n(g\circ l)$ est analytique sur l'ouvert Δ qui contient 0, donc $n=1$.

5.10. Proposition. *Soit f une fonction analytique, strictement injective sur le disque circonférencié $D(a, r^+)$ resp. non circonférencié $D(a, r^-)$. Alors $|f'(a)|=\alpha$ est non nulle et*

$$f(D(a, r^+)) = D(f(a), (\alpha r)^+)$$

$$f(D(a, r^-)) = D(f(a), (\alpha r)^-).$$

Démonstration. On se ramène au cas: $a=0$, $f(a)=0$, $f'(0)=1$ et $r=1$. Alors f satisfait aux hypothèses du lemme (5.2) (resp. (5.4)) avec $p=1$.

5.11. Proposition [on suppose que K est de caractéristique nulle]. *Toute fonction analytique injective est strictement injective.*

Démonstration. C'est une conséquence immédiate du théorème (5.5).

5.12. Lemme. *L'ensemble des transformations homographiques sans pôle dans A est une partie fermée de $\mathscr{F}(A)$.*

Démonstration. Soit $T_n(X)=\dfrac{a_n X+b_n}{c_n X+d_n}$ une suite d'homographies sans pôle dans A, convergeant uniformément sur A vers un élément analytique f. Quitte à prendre une sous-suite, on peut supposer que pour tout $n \in \mathbf{N}$, $T_n(X)$ est une P_n-fraction. Il en résulte que $d_n=1$, $|c_n|<1$, a_n et b_n appartiennent à A_{P_n} — et que pour tout $x \in A$ $f(x)-T_n(x) \in P_n$. Ceci implique que pour tout $x \in A$ $T_{n+1}(x)-T_n(x) \in P_n$ donc les éléments de $k_{P_n}(X)$,

$$\overline{T}_n(x) = \frac{\overline{a_n} X+\overline{b_n}}{\overline{c_n} X+\overline{1}} \quad \text{et} \quad \overline{T}_{n+1}(X) = \frac{\overline{a_{n+1}} X+\overline{b_{n+1}}}{\overline{c_{n+1}} X+\overline{1}}$$

sont égaux et par conséquent leurs coefficients, éléments de k_{p_n}, le sont. Donc, pour tout $n \in \mathbf{N}$, $a_n-a_{n+1} \in P_n$, $b_n-b_{n+1} \in P_n$ et $c_n-c_{n+1} \in P_n$. Les suites $(a_n)_{n\in\mathbf{N}}$, $(b_n)_{n\in\mathbf{N}}$ et $(c_n)_{n\in\mathbf{N}}$ sont des suites de Cauchy. Soient a, b, c leur limite respective, et soit T la transformation homographique

$$T(X) = \frac{aX+b}{cX+1}.$$

Pour tout $x \in A$ et tout $n \in \mathbf{N}$, $T(x)-T_n(x) \in P_n$, donc $T=\lim_{n\to\infty} T_n$ dans $H(A)$.

5.13. Proposition. Soit f une fonction analytique strictement injective sur A. On suppose que $f(0)=0$ et $f'(0)=1$. Alors f est une fonction homographique.

Démonstration. Soit $(P_n)_{n\in\mathbf{N}}$ une suite cofinale dans $\mathscr{P}$ et $(f_n)_{n\in\mathbf{N}}$ une suite approximante de f sur A. Il existe $N\in\mathbf{N}$ tel que pour tout $n>N$, f_n est une P_n-fraction,

$$f_n(x)-f(x)\in P_n \quad \text{et} \quad f_n'(x)+f'(x)\in P_n \quad \text{pour tout} \quad x\in A,$$

(d'après le corollaire (3.4)).

Notons $\overline{f_n}\colon k_{P_n}\cup\{\infty\}\to k_{P_n}\cup\{\infty\}$ la fonction définie par $\overline{f_n}(\bar{x})=\overline{f_n(x)}$ et montrons que pour n assez grand $\overline{f_n}$ est strictement injective sur $A/P_n=\{\bar{x}\in k_{p_n}\colon x\in A\}$. Pour n assez grand, et pour tout $x\in A$, on a $\bar{f}_n(\bar{x})=\overline{f(x)}$ et $\overline{f_n'}(\bar{x})=\overline{f'(x)}$ dans k_{P_n}. D'après la proposition 5.10, la fonction f étant strictement injective, on sait que $f(A)=A$ et, pour tout $x\in A$,

$$f(x+P_n)=f(x)+P_n \quad \text{et} \quad |f'(x)|=|f'(0)|=1.$$

Soient $a\in A$ et $b\in A$ tels que $\overline{f_n}(\bar{a})=\overline{f_n}(\bar{b})$, on a alors $f(a)-f(b)\in P_n$ et il existe $c\in P_n$ tel que $f(b)=f(a+c)$. Comme f est injective ceci entraîne que $a+c=b$ donc $b-a\in P_n$.

Ainsi la fonction $\overline{f_n}$ est injective sur A/P_n. De plus $|f_n'(x)|=1$ pour tout $x\in A$. Donc $\overline{f_n}$ est strictement injective sur A/P_n qui est une partie infinie de k_{P_n}. Donc la fraction rationnelle $\overline{f_n}(X)$ est homographique. Notons la

$$\overline{f_n}(X)=\frac{\overline{\alpha_n}\,X+\overline{\beta_n}}{\overline{\gamma_n}\,X+\overline{\sigma_n}}$$

et soit $T_n(X)\in K(X)$ la fraction rationnelle définie par

$$T_n(X)=\frac{\alpha_n X+\beta_n}{\gamma_n X+\sigma_n}.$$

Pour tout $x\in A$, on a

$$T_n(x)-f(x)=\big(T_n(x)-f_n(x)\big)+\big(f_n(x)-f(x)\big)\in P_n$$

et par conséquent pour tout $n>N$

$$T_n(x)-T_{n+1}(x)\in P_n.$$

On en déduit que la suite $(T_n)_{n\in\mathbf{N}}$ converge uniformément sur A et d'après le lemme 5.12 sa limite est une homographie.

5.14 Théorème. *Toute fonction analytique strictement injective sur un ouvert de K est une fonction homographique.*

Démonstration. Soit $D(a,r^+)$ un disque contenu dans D. La fonction $F(u)=\frac{1}{rf'(a)}[f(a+ru)-f(a)]$ vérifie les hypothèses de la proposition précédente. C'est une fonction homographique, donc la restriction de f à $D(a,r^+)$ est une fonction homographique et d'après l'unicité du prolongement analytique f est une fonction homographique sur D.

5.15. Corollaire. *Soit D un ouvert de K. Toute fonction f: $D \to f(D)$ analytique et strictement injective est une fonction bianalytique. Si K est de caractéristique nulle, toute fonction f: $D \to f(D)$ analytique et injective est bianalytique.*

5.16. Corollaire. *Soit f une fonction analytique strictement injective sur un P-ouvert D. Il existe un idéal premier $Q \in \mathscr{P}$ tel que $f(D)$ soit un Q-ouvert.*

5.17. Remarques. Ces résultats concernant les fonctions bianalytiques dans les corps hédériques sont très différents de ceux que l'on a en analyse complexe et en analyse ultramétrique «de rang un». Rappelons que sur les corps valués de rang un le problème des fonctions bianalytiques est loin d'être résolu (voir [1] et [9]). Dans [1] A. Escassut a montré que si D est une partie ultracirconférenciée d'un corps valué de caractéristique nulle tout élément analytique f sur D, injectif est un élément bianalytique de D sur $f(D)$. E. Motzkin a obtenu un résultat analogue pour les quasi-connexes n'ayant qu'un nombre fini de trous ouverts [9].

VI. Transformations conformes et isomorphismes d'algèbres de fonctions analytiques

6.1. Rappelons que deux ouverts D et Δ de K sont dits *conformément équivalents* s'il existe une application bianalytique h de D sur Δ. S'il existe une application bianalytique h de l'ouvert D sur l'ouvert Δ, l'application ψ: $\mathscr{F}(\Delta) \to \mathscr{F}(D)$ définie par $\psi(f) = f \circ h$ pour tout $f \in \mathscr{F}(\Delta)$ est un K-isomorphisme d'algèbres. On se propose d'établir ici la réciproque de ce résultat: Les algèbres $\mathscr{F}(D)$ et $\mathscr{F}(\Delta)$ étant K-isomorphes les ouverts D et Δ sont conformément équivalents.

K-homomorphismes d'algèbres de fonctions analytiques

6.2. Lemme. *Soient D et Δ deux ouverts quelconques de K et ψ: $\mathscr{F}(\Delta) \to \mathscr{F}(D)$ un K-homomorphisme d'algèbres $(\psi(1)=1)$. On pose $h = \psi(Id_\Delta)$, on a alors*

i) $h(D) \subset \Delta$;

ii) *si ψ est un K-isomorphisme* $h(D) = \Delta$;

iii) *pour tout* $f \in \mathscr{F}(\Delta)$, $\psi(f) = f \circ h$.

Démonstration. i) Soit $a \in D$. Si $h(a) \notin \Delta$, alors $Id_\Delta - h(a)$ est inversible dans $\mathscr{F}(\Delta)$. Soit g son. Ainsi $g(Id_\Delta - h(a)) = 1$ donc $\psi(g) \cdot (h - h(a)) = 1$ ce qui est impossible puisque $h - h(a)$ s'annule au point a.

ii) Supposons que ψ soit un K-isomorphisme. Soit $b \in \Delta$. Si $b \notin h(D)$ alors $h - b$ est inversible dans $\mathscr{F}(D)$, donc $\psi^{-1}(h-b) = Id_\Delta - b$ est inversible dans $\mathscr{F}(\Delta)$, ce qui est faux.

iii) Soient $a \in D$ et $b = h(a)$. Pour chaque $f \in \mathscr{F}(\Delta)$, posons

$$\eta(f) = \psi(f)(a)$$

$$\xi(f) = f \circ h(a).$$

Ainsi η et ξ sont des homomorphismes surjectifs de l'algèbre $\mathscr{F}(D)$ dans K. Le noyau de ξ est $\xi^{-1}(0) = \{f \in \mathscr{F}(\Delta) : f(b) = 0\} = (Id_\Delta - b)$. $\mathscr{F}(\Delta)$ Et $\xi^{-1}(0) \subset \eta^{-1}(0)$

donc les noyaux de ξ et de η sont égaux. La fonction $f-f(b)\in\xi^{-1}(0)$ donc $f-f(b)\in \in\eta^{-1}(0)$, c'est-à-dire $\psi(f)(a)=f(b)=f[h(a)]$.

6.3. LEMME. *Si D et Δ sont deux P-ouverts, tout K-homomorphisme de $\mathscr{F}(\Delta)$ dans $\mathscr{F}(D)$ est continu.*

En effet, soit h: $D\to\Delta$ une fonction analytique sur D. Pour tout $\alpha\in\Gamma$, $f\circ h\in \underset{\mathscr{F}(D)}{V}(0,\alpha)$ dès que $f\in \underset{\mathscr{F}(\Delta)}{V}(0,\alpha)$.

6.4. LEMME. *Soient D et Δ deux ouverts quelconques, h: $D\to\Delta$ une fonction analytique strictement injective. Alors le K-homomorphisme d'algèbres ξ: $\mathscr{F}(\Delta)\to \to\mathscr{F}(D)$ défini par $\xi(f)=f\circ h$ est continu.*

DÉMONSTRATION. Démontrons la continuité en 0. Soit $V(0, D(P), P)$ un voisinage de 0 dans $\mathscr{F}(D)$. D'après la proposition (5.16) il existe un idéal premier R de A tel que $h(D(P))$ soit un R-ouvert, donc $h(D(P))$ est contenu dans $\Delta(R)$. Soit $Q\subset R\cap P$. On a $\Delta(Q)\supset\Delta(R)$. Si $f\in V(0, \Delta(Q), Q)$, c'est-à-dire si $f(y)\in Q$ pour tout $y\in\Delta(Q)$ alors $f\circ h(x)\in Q\subset P$ pour tout $x\in D(P)$; c'est-à-dire $f\circ h\in V(0, D(P)\, P)$.

6.5. LEMME. *Soient D et Δ deux ouverts quelconques de K, ψ un K-homomorphisme d'algèbres de $\mathscr{F}(\Delta)$ dans $\mathscr{F}(D)$. Ou bien il existe $a\in D$ tel que $\psi(f)=f(a)$ pour tout $f\in\mathscr{F}(\Delta)$ ou bien ψ est injectif.*

DÉMONSTRATION. Soit $h\in\mathscr{F}(D)$ telle que $\psi(f)=f\circ h$. Ou bien h est constante ou bien h est une application ouverte (théorème (5.3)). Si h est constante: $h(x)=a$ alors $\psi(f)=f(a)$. Si h est ouverte $h(D)$ est un ouvert de Δ, et toute fonction analytique f qui est nulle sur $h(D)$ est nulle sur Δ tout entier donc $\psi(f)=0$ implique $f=0$.

6.6. THÉORÈME. *Soit D et Δ des ouverts quelconques de K et ψ: $\mathscr{F}(\Delta)\to\mathscr{F}(D)$ un K-isomorphisme d'algèbres. Alors D et Δ sont conformément équivalents et ψ est continu. Plus précisément, soient*

$$h=\psi(id_\Delta);\quad k=\psi^{-1}(id_D).$$

Alors h: $D\to\Delta$ et k: $\Delta\to D$ sont deux fonctions homographiques réciproques l'une de l'autre.

DÉMONSTRATION. a) h et k sont des fonctions analytiques.

b) $h(D)=id_\Delta(\Delta)=\Delta$

$k(\Delta)=id_D(D)=D$

d'après le lemme (6.2).

c) Soit $x_0\in D$ la fonction analytique $h-h(x_0)$ s'annule sur D, donc (d'après la proposition 1.3), il existe $g\in\mathscr{F}(D)$ telle que

$$h(x)-h(x_0)=(x-x_0)\,g(x)\quad\text{pour tout}\quad x\in D.$$

On va montrer que g est inversible dans $\mathscr{F}(D)$.

Supposons le contraire, alors il existe $x_1 \in D$ et $l \in \mathscr{F}(D)$ tels que

$$g(x) = (x - x_1)\, l(x)$$

donc

$$h(x) - h(x_0) = (x - x_0)(x - x_1)\, l(x)$$

c'est-à-dire

$$h - h(x_0) = (id_\Delta - x_0)(id_\Delta - x_1)\, l.$$

En prenant l'image par ψ^{-1} des deux membres il vient

$$id_\Delta - h(x_0) = (k - x_0)(h - x_1)\psi^{-1}(l).$$

Or $id_\Delta - h(x_0)$ est analytique sur Δ, $k - x_0$ et $h - x_1$ s'annulent sur Δ, donc $\psi^{-1}(l)$ ne serait pas analytique sur Δ. La fonction g est donc inversible dans $\mathscr{F}(D)$.

Prenons l'image par ψ^{-1} des deux membres de l'égalité

$$h - h(x_0) = (id_\Delta - x_0)\, g.$$

On obtient

$$id_\Delta - h(x_0) = (k - x_0)\eta \quad \text{où} \quad \eta = \psi^{-1}(g) \quad \text{et} \quad \frac{1}{\eta} \in \mathscr{F}(\Delta).$$

Donc $k - x_0 = [id_\Delta - h(x_0)]\dfrac{1}{\eta}$ de sorte que

$$k\,[h(x_0)] - x = [h(x_0 - h(x_0)]\frac{1}{\eta}(h(x_0)) = 0.$$

Donc $k[h(x_0)] = x_0$ pour tout $x_0 \in D$. Autrement dit $k \circ h = id_D$. De même $h \circ k = id_\Delta$. Les fonctions analytiques h et k sont strictement injectives d'après la proposition (5.9), ce sont donc des fonctions homographiques (théorème 5.14). Et l'on a

$$\psi(f) = f \circ h \quad \text{pour tout} \quad f \in \mathscr{F}(D),$$

$$\psi^{-1}(g) = g \circ h \quad \text{pour tout} \quad g \in \mathscr{F}(\Delta),$$

donc ψ et ψ^{-1} sont continues.

Compte tenu des résultats précédents il est facile de caractériser les ouverts Δ conformément équivalents à un ouvert D de K, ainsi que les K-isomorphismes d'algèbres de fonctions analytiques.

6.7. Théorème. *Pour que deux ouverts D et Δ soient conformément équivalents il faut et il suffit qu'il existe une transformation homographique T telle que $T(D) = \Delta$. Pour qu'il existe un K-isomorphisme d'algèbres $\psi : \mathscr{F}(\Delta) \to \mathscr{F}(D)$ il faut et il suffit que D et Δ soient conformément équivalents et s'il en est ainsi il existe alors une transformation homographique T telle que $\psi(f) = f \circ T$ pour tout $f \in \mathscr{F}(\Delta)$.*

RÉFÉRENCES

[1] ESCASSUT, A., Éléments spectralement injectifs et générateurs universels dans une algèbre de Tate, *Collect. Math.* **28** (1977), 131—148. *MR* **58** # 22658.

[2] FENEYROL-PERRIN, Y., Fonctions analytiques d'une où de plusieurs variables sur certains corps values au sens de Krull, *Thèse de Doctorat d'Etat*, Université de Paris VI, 1980; *C. R. Acad. Sci. Paris Sér. A—B* **290** (1980), A209—A211. *MR* **81a**: 12019.

[3] FENEYROL-PERRIN, Y., Répartition des zéros des éléments et des fonctions analytiques sur un ouvert d'un corps valué au sens de Krull, complet et algébriquement clos, *C. R. Acad. Sci. Paris Sér. A—B* **287** (1978), A131—A132. *MR* **58** # 10834.

[4] FENEYROL-PERRIN, Y., Fonctions analytiques dans les corps valués de rang supérieur à un, *Compositio Math.* **49** (1983), 51—74.

[5] FENEYROL-PERRIN, Y. et HADDAD, L., A non-Archimedean Mittag—Leffler theorem, *C. R. Math. Rep. Acad. Sci. Canada* **6** (1984), 199—204 (Not in *MR*).

[6] FENEYROL-PERRIN, Y. et HADDAD, L., Un théorème de Mittag—Leffler pour les fonctions méromorphes sur un corps valué au sens de Krull, *Compositio Math.* **57** (1986), 249—269.

[7] HOBEIKA, R., Séries de Taylor et séries de Laurent dans certains corps valués au sens de Krull, *Thèse de 3ème cycle,* Université de Paris VI, 1976.

[8] KRASNER, M., Rapport sur le prolongement analytique dans les corps valués complets par la méthode des éléments analytiques quasi-connexes, *Bull. Soc. Math. France,* Mémoire No. 39—40, Soc. Math. France, Paris, 1974, 131—254. *MR* **52** # 6033.

[9] MOTZKIN, E., L'arbre d'un quasi-connexe: un invariant conforme p-adique, *Groupe d'étude d'analyse ultramétrique,* 3ème année, 1981—1982, I.H.P., Paris.

(Reçu le 6 août 1986)

UNIVERSITÉ DE CLERMONT
COMPLEXE SCIENTIFIQUE DES CÉZEAUX
MATHÉMATIQUES PURES
B. P. 45
F—63170 AUBIÈRE
FRANCE

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 239—258

HYPERSOLVABLE AND SUPERNILPOTENT RADICALS OF NEAR-RINGS

R. MLITZ and A. OSWALD

For associative rings, the supernilpotent radicals have been shown by Andrunakievich [2], resp. Rjabuhin [12], to have some particularly nice properties. For example, there is a formula (involving two-sided Noetherian quotients) for the computation of the radical of a ring A knowing the radicals of an ideal I and of A/I; moreover, these radicals are exactly those whose semisimple classes are the subdirect closures of weakly special classes.

In the present paper, the ideas of Andrunakievich and Rjabuhin are investigated in the near-ring case. In Section 2 the hypersolvable radicals (i.e. those for which all nilpotent near-rings are radical) are characterized by formulae involving Noetherian quotients. In Section 3 we obtain the near-ring analogues of Andrunakievich's formula for supernilpotent radicals (i.e. hypersolvable radicals with hereditary radical class). In both sections, we exhibit some hereditary, respectively nearly hereditary, properties of semisimple classes. In the last section, we introduce a near-ring analogue to Rjabuhin's weakly special classes. If either the radical class or the semisimple class consists only of zero-symmetric near-rings, we show that the radical is supernilpotent if and only if the semisimple class is the subdirect closure of a weakly special class. Since near-rings have only one-sided distributivity, we mostly use one-sided Noetherian quotients. In doing so we obtain results which even in the ring case are new. Notice that because we use one-sided quotients and because semisimple classes are not, in general, hereditary the proof methods used for rings do not carry over to the near-ring case.

1. Definitions and basic results

The book by Pilz [11] provides an excellent introduction to the theory of near-rings. However, for the sake of completeness we begin with a brief summary of the concepts and results which we will need.

A *near-ring* is a triple $(N, +, \cdot)$ such that $(N, +)$ is a (not necessarily abelian) group with identity 0, $(N, \cdot)$ is a semigroup and for all $n, m, p \in N$, $(n+m)p = np+mp$. (Strictly speaking this is a right near-ring.) If N is a near-ring and for all $n \in N$, $n0=0$ then N is called *zero-symmetric* whilst if $nm=n$ for all $n, m \in N$ then N is called *constant*. Every near-ring N can be written as the sum of its zero symmetric part N_0 and its constant part N_C.

1980 *Mathematics Subject Classification* (1985 Revision). Primary 16A76.
Key words and phrases. Radical, near-ring.

We write $N^k=\{n_1n_2\dots n_k|n_i\in N\}$. Then N is a *zero near-ring* if $N^2=0$ and, in general, N is *nilpotent* if $N^k=0$, for some positive integer k.

We denote the additive group of N by N^+ and write $K^+\leqq N^+$ to indicate that the subset K of N is a subgroup of N^+ and similarly $K^+\lhd N^+$ if K^+ is a normal subgroup of N^+. If $K^+\leqq N^+$ and $NK\subseteqq K$ we say that K is an *N-subgroup of N* and write $K\underset{N}{\leqq}N$. When $K\underset{N}{\lhd}N$ we have a *normal N-subgroup of N*. In the case where $K\underset{N}{\leqq}N$ and $KN\subseteqq K$ we call K an *invariant subgroup* of N (and similarly for $K\underset{N}{\lhd}N$) and write $K\ll N$ ($K\ll\!\lhd N$).

A subset K of N is called *a right ideal* of N, written $K\underset{r}{\lhd}N$ if $K^+\lhd N^+$ and $KN\subseteqq K$. If $K^+\lhd N^+$, and for all $m,n\in N$, $k\in K$ we have $m(n+k)-mn\in K$ then K is called *a left ideal* of N and we write $K\underset{l}{\lhd}N$.

A subset K of N which is both a left and a right ideal is called an *ideal* and we use the notation $K\lhd N$. In the usual way, near-ring homomorphisms are defined and there is a 1—1 correspondence between ideals of N and kernels of near-ring homomorphisms with domain N. If K is an ideal of N we denote the corresponding homomorphism by φ_K. Then $\varphi_K\colon N\to N/K$ as usual.

A class $\mathfrak{C}$ of near-rings is *hereditary with respect to* N-subgroups (normal N-subgroups, etc.) if $N\in\mathfrak{C}$ and $K\underset{N}{\leqq}N$ ($K\underset{N}{\lhd}N$, etc.) imply $K\in\mathfrak{C}$; hereditary with respect to ideals will simply be called *hereditary*. If $N\in\mathfrak{C}$ implies $\varphi N\in\mathfrak{C}$ for every near-ring homomorphism φ with domain N then $\mathfrak{C}$ is *homomorphically closed.* A class $\mathfrak{U}$ of near-rings which is hereditary and homomorphically closed is a *universal class.*

Turning now to radical theory, a radical in the sense of Hoehnke (an *H-radical*) on a universal class $\mathfrak{U}$ of near-rings is a mapping R assigning to each $N\in\mathfrak{U}$, an ideal $R(N)$ of N such that

1. $\varphi R(N)\subseteqq R(\varphi N)$ for every near-ring homomorphism φ with domain N and
2. $R(N/R(N))=0$.

If in addition

3. $R(R(N))=R(N)$ and
4. $R(I)=I$ for $I\lhd N$ implies $I\subseteqq R(N)$

then R is a radical in the sense of Kurosh and Amitsur (a KA-radical). From Mlitz [8] this definition coincides with the classical one. If R is an H-radical or a KA-radical on a universal class $\mathfrak{U}$ of near-rings we denote by $\mathbf{R}$ the corresponding *radical class* (i.e. the class of all near-rings N with $R(N)=N$) and by $\mathbf{S}$ the corresponding *semi-simple class* (the class of all near-rings N with $R(N)=0$).

If T and U are subsets of a near-ring N we define
the *left Noetherian quotient*, $(T\colon U)$, by $(T\colon U)=\{n\in N|nU\subseteqq T\}$
the *right Noetherian quotient*, $(T\colon U)'$, by $(T\colon U)'=\{n\in N|Un\subseteqq T\}$
the *two-sided Noetherian quotient*, $(T\colon U)''$, by

$$(T\colon U)''=\{n\in N|nU\subseteqq T \quad\text{and}\quad Un\subseteqq T\}.$$

Observe that $(T\colon U)''=(T\colon U)\cap(T\colon U)'$. If M is a subnear-ring of N we write $(T\colon U)\cap M$ etc. for the corresponding quotients in M.

If L is a subset of N we write L_N for the maximal ideal of N and L_N^l for the maximal left ideal of N contained in L (if they exist). Furthermore, $\langle L\rangle$ will denote the ideal of N generated by L.

We now have the following basic results, the first of which follows from Wiegandt [14] and the others from the definitions and by straightforward calculation.

1.1. A subclass $\mathbf{R}$ of a universal class $\mathfrak{U}$ of near-rings in a KA-radical class if and only if

(i) $\mathbf{R}$ is homomorphically closed;

(ii) $\mathbf{R}$ is closed with respect to sums of ideals of the same near-ring;

(iii) $\mathbf{R}$ is closed with respect to extensions in $\mathfrak{U}$, i.e. if $N\in\mathfrak{U}$ and $N/K\in\mathbf{R}$, $K\in\mathbf{R}$ then $N\in\mathbf{R}$.

1.2. Every radical R on a universal class $\mathfrak{U}$ of near-rings can be represented in the form

$$R(N)=\varphi_I^{-1}\,R(N/I)\cap K(I)$$

where $K(I)$ is a suitable subset of N depending upon $I\triangleleft N$.

1.3. If $I\triangleleft N$ and $N_c\subseteqq I$ then $I\mathrel{<\!\triangleleft} N$.

1.4. If $T^+\leqq N^+$ then $(T\colon U)_N$ exists.

1.5. If $U\underset{N}{\leqq}N$ and $T^+\triangleleft U^+$ =then $(T\colon U)\underset{r}{\triangleleft}N$.

1.6. If $T\underset{l}{\triangleleft}N$ and $NU\subseteqq U$ then $(T\colon U)\triangleleft N$.

1.7. If $U\leqq N$ and $T\underset{r}{\triangleleft}N$ then $T\subseteqq(T\colon U)$.

1.8. If $U\leqq N$ and $T\underset{l}{\triangleleft}N$ then $UT\subseteqq T+U_c$; in particular, if $U_c\subseteqq T$ then $T\subseteqq(T\colon U)'$.

1.9. If $I\triangleleft N$, $T\subseteqq N$ then, provided T_N exists, we have $(T\cap I)_N=T_N\cap I$.

2. Hypersolvability

A radical R on a universal class $\mathfrak{U}$ of near-rings is *hypersolvable* if $R(N)=N$ for every nilpotent near-ring N in $\mathfrak{U}$. Our first result seems to be well-known but to the best of our knowledge, no proof has ever been published. Hence for completeness we have

PROPOSITION 2.1. *A KA-radical R on a universal class $\mathfrak{U}$ of near-rings is hypersolvable if and only if the corresponding radical class $\mathbf{R}$ contains all zero near-rings belonging to $\mathfrak{U}$.*

PROOF. Suppose $N^k=0$ and $N^{k-1}\neq 0$ and let $I=\langle N^{k-1}\rangle$. Then $N^{k-1}I\subseteqq\subseteqq N^k=0$ and so $N^{k-1}\subseteqq(0\colon I)$. From 1.6, $(0\colon I)$ is an ideal of N and so $\langle N^{k-1}\rangle\subseteqq\subseteqq(0\colon I)$, i.e. $I\subseteqq(0\colon I)$. It follows that $I^2=0$. Now suppose that every zero near-ring in $\mathfrak{U}$ is in $\mathbf{R}$ and let $N\in\mathfrak{U}$ be nilpotent. Then $\bar{N}=N/R(N)$ is also nilpotent. If $\bar{N}=\bar{0}$ then $N=R(N)$ and $N\in\mathbf{R}$. If $\bar{N}^k=\bar{0}$ and $\bar{N}^{k-1}\neq\bar{0}$, for some $k>1$, then with $\bar{I}=\langle\bar{N}^{k-1}\rangle$ we have $\bar{I}^2=\bar{0}$. Now if $I\lhd N$ with $\bar{I}=I/R(N)$, then $I\in\mathfrak{U}$ and so $\bar{I}\in\mathfrak{U}$ and hence $R(\bar{I})=\bar{I}$. It follows that $\bar{I}\subseteq R(N/R(N))=\bar{0}$. This contradicts $\bar{N}^{k-1}\neq\bar{0}$ and so $N=R(N)$ as required and R is hypersolvable. Conversely, if R is hypersolvable then every nilpotent near-ring is in $\mathbf{R}$ and so every zero near-ring is in $\mathbf{R}$.

Our main aim will be to characterize hypersolvability via Noetherian quotients and for this purpose we need

Proposition 2.2. *If R is a hypersolvable KA-radical and I and K are ideals of N with $I^2\subseteq K$ and $N/K\in\mathbf{S}$ then $I\subseteq K$.*

Proof. Since $I^2\subseteq K$, $\varphi_K(I^2)=0$ and so $\varphi_K(I)^2=0$. It follows that $\varphi_K(I)\in\mathbf{R}$. But $\varphi_K(N)\in\mathbf{S}$ and so $\varphi_K(I)=0$ and $I\subseteq K$ as required.

Theorem 2.3. *For a KA-radical R on a universal class $\mathfrak{U}$ of near-rings the following are equivalent:*

(i) *R is hypersolvable;*
(ii) *$K=(K\colon N)$ for all $N\in\mathfrak{U}$ and $K\lhd N$ with $N/K\in\mathbf{S}$;*
(iii) *$R(N)=(R(N)\colon N)$ for all $N\in\mathfrak{U}$;*
(iv) *$R(N)=\varphi_I^{-1}R(N/I)\cap(R(N)\colon N)$ for all $N\in\mathfrak{U}$, $I\lhd N$;*
(v) *$R(N)=[\varphi_I^{-1}R(N/I)\cap(R(N)\colon U)]_N$ for all $N\in\mathfrak{U}$, $I\lhd N$ and all subsets U of N containing $\varphi_I^{-1}R(N/I)$;*
(vi) *$R(N)=[\varphi_I^{-1}R(N/I)\cap(T\colon U)]_N$ for all $N\in\mathfrak{U}$, $I\lhd N$ and suitable T, $U\subseteq N$ (which may depend on I);*
(vii) *$(O\colon N)\subseteq R(N)$ for all $N\in\mathfrak{U}$;*
(viii) *$\varphi_I^{-1}R(N/I)\cap(O\colon N)\subseteq R(N)$ for all $N\in\mathfrak{U}$, $I\lhd N$.*

Proof. (i)⇒(ii). Let $N\in\mathfrak{U}$ and $K\lhd N$ with $N/K\in\mathbf{S}$. From 1.7, $K\subseteq(K\colon N)$. Now $(K\colon N)^2\subseteq(K\colon N)N\subseteq K$ and so by 2.2, $(K\colon N)\subseteq K$. It follows that $K=(K\colon N)$.

(ii)⇒(iii)⇒(vii)⇒(viii) are trivial as also are (iv)⇒(vi), (v)⇒(vi) and (iii)⇒(iv).

(i)⇒(v). From the definition of a radical $\varphi_I^{-1}R(N/I)\subseteq U$ implies $R(N)\subseteq U$ and since $R(N)\lhd N$, $R(N)\subseteq(R(N)\colon U)$. Hence $R(N)\subseteq[\varphi_I^{-1}R(N/I)\cap(R(N)\colon U)]_N$. If $U=\varphi_I^{-1}R(N/I)$ then $[U\cap(R(N)\colon U)]_N=U\cap(R(N)\colon U)$ and $[U\cap(R(N)\colon U)]^2\subseteq(R(N)\colon U)U\subseteq R(N)$ so that by 2.2 $U\cap(R(N)\colon U)\subseteq R(N)$. Alternatively, if $\varphi_I^{-1}R(N/I)\subseteq U$, then $(R(N)\colon N)\subseteq(R(N)\colon U)\subseteq(R(N)\colon\varphi_I^{-1}R(N/I))$. Since (i)⇒⇒(iv) it follows that $R(N)=\varphi_I^{-1}R(N/I))\cap(R(N)\colon N)\subseteq[\varphi_I^{-1}R(N/I)\cap(R(N)\colon U)]_N\subseteq\varphi_I^{-1}R(N/I)\cap(R(N)\colon\varphi_I^{-1}R(N/I))\subseteq R(N)$ and hence we have (i)⇒(v).

(vi)⇒(viii). For suitable $T, U, R(N)\subseteq(T\colon U)$. Hence $0=0\cdot U\subseteq R(N)U\subseteq T$ and so $(O\colon N)\subseteq(T\colon U)$. But $(O\colon N)\lhd N$ and so we have (viii).

Finally, (viii)⇒(i). If $N^2=0$ then applying (viii) with $I=N$ gives $N\cap(O\colon N)\subseteq\subseteq R(N)$ and thus $N=R(N)$ and R is hypersolvable.

(Note that in proving (i)⇒(v) we also proved that for any subset U of N containing $\varphi_I^{-1}R(N/I)$, the set $\varphi_I^{-1}R(N/I)\cap(R(N)\colon U)$ is actually an ideal of N.)

Corollary 2.4. *If R is a hypersolvable radical on a universal class $\mathfrak{U}$ of near-rings and if $N, G\in\mathfrak{U}$ with G a normal invariant N-subgroup of N then $R(G)$ is a right ideal of N. In particular, if $I\lhd N$ and $N_c\subseteq I$ then $R(I)\lhd_r N$.*

PROOF. By 2.3 (iii), $R(G)=(R(G)\colon G)$ and so, in N, $R(G)=(R(G)\colon G)\cap G$. Since $G\underset{N}{\leqq}N$ and $R(G)^+\lhd G^+$ we have by 1.5 $(R(G)\colon G)\underset{r}{\lhd}N$. Furthermore, $G\underset{r}{\lhd}N$ and it then follows that $R(G)=(R(G)\colon G)\cap G\underset{r}{\lhd}N$ as required.

For the particular case of $I\lhd N$ with $N_c\subseteqq I$ we appeal to 1.3 and this general result to obtain $R(I)\underset{r}{\lhd}N$.

We say that a radical on $\mathfrak{U}$ is *right-strong* if for each $N\in\mathfrak{U}$, all radical right ideals of N are contained in $R(N)$. Right-strong radicals on the universal class of all zero-symmetric near-rings include the important J_2 and J_3 (Anderson—Kaarli—Wiegandt [1]).

COROLLARY 2.5. *If a hypersolvable radical on a universal class is right-strong then the corresponding semisimple class* **S** *is hereditary with respect to normal invariant subgroups belonging to* $\mathfrak{U}$.

In particular, the semisimple classes of the right-strong supernilpotent radicals are hereditary with respect to ideals containing the constant part of the near-ring under consideration. Because J_2 and J_3 are right-strong hypersolvable radicals, this result provides a new proof of the well-known hereditary nature of the corresponding semisimple classes.

Our final corollary uses the following result of Betsch and Kaarli [3]: If the semisimple class of a KA-radical of near-rings is hereditary, then either the radical is hypersolvable or it is trivial (i.e. **R** consists of one-element near-rings). Their work is in the universal class $\mathfrak{U}$ of all near-rings. It is straightforward to see that their results rely only on the fact that $\mathfrak{U}$ contains certain near-rings. They hold for example, if $\mathfrak{U}$ satisfies the following property

(P): For all $A, B\in\mathfrak{U}$, every near-ring whose additive group is $A^+\oplus B^+$ is in $\mathfrak{U}$.

COROLLARY 2.6. *If the universal class* $\mathfrak{U}$ *has property* (P), *then every right-strong KA-radical on* $\mathfrak{U}$ *whose semisimple class consists only of zero-symmetric near-rings has a hereditary semisimple class.*

We should mention that assuming some restrictions on the universal class $\mathfrak{U}$, the result of 2.5 has been obtained by Anderson, Kaarli and Wiegandt [1] using different methods.

In Theorem 2.3, the right-hand sides of the formulae in (iii), (iv) and (v) are heavily dependent upon $R(N)$. In general this will reduce their value as tools for the computation of $R(N)$ from a knowledge of $R(N/I)$. However, provided reasonable conditions on T and U can be found the formula in 2.3 (vi) will enable us to recover $R(N)$ from $R(N/I)$.

PROPOSITION 2.7. *Let R be a KA-radical on a universal class $\mathfrak{U}$ of near-rings and $N\in\mathfrak{U}$. For each triple (I, U, T) with $I\lhd N$ $T, U\subset N$ let*

$$X(I, U, T)=(\varphi_I^{-1}R(N/I)\cap(T\colon U))_N.$$

Then the following are equivalent:

(i) $R(N)=X(I, U, T)$;

(ii) (C1) $R(N)\cdot U \subseteq T$

(C2) $K \subseteq R(N)$ *and* $X(I, U, T)/K \in \mathbf{R}$ *for every ideal* K *of* $X(I, U, T)$;

(iii) (C1) $R(N)\cdot U \subseteq T$,

(C3) $K \subseteq R(N)$ *and* $X(I, U, T)/K \in \mathbf{R}$ *for some ideal* K *of* $X(I, U, T)$.

PROOF. (i)⇒(ii). Since $R(N)=X(I, U, T)$, $R(N)\subseteq(T\colon U)$ and so $R(N)\cdot U\subseteq T$ which is (C1). Since $\mathbf{R}$ is homomorphically closed (C2) follows immediately.

(ii)⇒(iii). This is trivial.

(iii)⇒(i). From (C1), $R(N)\subseteq(T\colon U)$ and so $R(N)\subseteq X(I, U, T)$. Now

$$X(I, U, T)/R(N) \simeq \frac{X(I, U, T)/K}{R(N)/K} \in \mathbf{R}.$$

It follows that $X(I, U, T)\in \mathbf{R}$ and so $X(I, U, T)\subseteq R(N)$. Hence $R(N)=X(I, U, T)$ as required.

In connection with this result we observe the following

REMARK 2.8. From 1.4 and 1.9 we see that for $T^+ \leqq N^+$ the ideal $X(I, U, T)$ simplifies to

$$X(I, U, T) = \varphi_I^{-1} R(N/I) \cap (T\colon U)_N.$$

Furthermore, if $T \lhd N$ and $NU \subseteq U$ when using 1.6 we get

$$X(I, U, T) = \varphi_I^{-1} R(N/I) \cap (T\colon U).$$

In particular, using 2.3 (v) we see that for all $I \lhd N$,

$$R(N) = \varphi_I^{-1} R(N/I) \cap \big(R(N)\colon \varphi_I^{-1} R(N/I) + N_C\big)$$

when R is hypersolvable.

REMARK 2.9. In the case where R is hypersolvable we see from 2.3 (v) that for any ideal $I \lhd N$ there are pairs (T, U) satisfying the conditions of 2.7 whilst when R is not hypersolvable there is at least one ideal $I \lhd N$ for which no such pair exists.

PROPOSITION 2.10. *Let R be hypersolvable and $U \subseteq \varphi_I^{-1} R(N/I)$. Sufficient conditions for $R(N)=X(I, U, T)$ are*

(C1) *and*

(C4) *there is an ideal J of N with $J \subseteq U$ and*

$$T \cap J \subseteq R(N), \quad X(I, U, T)/(J \cap X(I, U, T)) \in \mathbf{R}.$$

PROOF. Since $J \subseteq U$, $J \cap (T\colon U) \subseteq U$ and so

$$(J\cap(T\colon U))^2 \subseteq (J\cap(T\colon U))\, U \subseteq J\cap(T\colon U)\, U \subseteq J \cap T \subseteq R(N).$$

Then by 1.9 and 2.2,

$$\begin{aligned} J\cap X(I, U, T) &= J\cap\big(\varphi_I^{-1} R(N/I)\cap(T\colon U)\big)_N \\ &= \big(J\cap\varphi_I^{-1} R(N/I)\cap(T\colon U)\big)_N \\ &= \big(J\cap(T\colon U)\big)_N \\ &\subseteq R(N). \end{aligned}$$

Hence we obtain (C3) with $K=J\cap X(I, U, T)$ and the result follows from 2.7.

PROPOSITION 2.11. *Let R be hypersolvable and U be a subnear-ring of N containing $\varphi_I^{-1}R(N/I)$. Sufficient conditions for $R(N)=X(I, U, T)$ are*
(C1) *and*
(C5) $T\cap U \lhd U$, $T\cap X(I, U, T)\in \mathbf{R}$.

PROOF. Since $\varphi_I^{-1}R(N/I)\subseteq U$ and $X(I, U, T)\subseteq \varphi_I^{-1}R(N/I)$, $T\cap X(I, U, T)= =(T\cap U)\cap X(I, U, T)$ and so $T\cap X(I, U, T)$ is an ideal of $X(I, U, T)$ by (C5). But $(X(I, U, T))^2\subseteq (T\colon U)\cdot U\cap X(I, U, T)\subseteq T\cap X(I, U, T)$ and so

$$X(I, U, T)/\,(T\cap X(I, U, T))$$

is a zero near-ring and hence in $\mathbf{R}$.

Since $\mathbf{R}$ has the extension property, $X(I, U, T)\in \mathbf{R}$. Applying 2.7 (iii) with $K=(0)$ we have the result.

Observe that when $\varphi_I^{-1}R(N/I)\subseteq U$, the condition (C1) will be satisfied if $R(U)\cdot U\subseteq T$ since $R(N)\subseteq R(U)$. It follows that the conditions in 2.11 can be replaced by
(C1) $R(U)\cdot U\subseteq T$ and
(C5).
These conditions are, of course, stronger than those of (2.11). However, because they do not involve $R(N)$ they might be more useful in the computation of $R(N)$.

When the universal class $\mathfrak{U}$ consists only of rings or only of distributive near-rings, these results carry over from left quotients to right quotients by symmetry and then to two-sided quotients by intersection. In the general case we cannot expect such a symmetry; however, under some restrictions (which will automatically be satisfied by zero-symmetric near-rings) we can obtain results both for right and for two-sided quotients which are similar to those for left quotients. Our next theorem is analogous to 2.3. Rather than rewrite two modified versions of that theorem we indicate the alterations to its equivalent conditions as follows:

We obtain $(j)'$, $(j)''$ from (j) in 2.3 by
$j = 2, 3, 4$ – replace the left quotient $(A\colon B)$ appearing in (j) by $(A\colon B)'_N$, respectively $(A\colon B)''$;
$j = 5, 6$ – replace the left quotient $(A\colon B)$ appearing in (j) by $(A\colon B)'$, respectively $(A\colon B)''$;
$j = 7, 8$ – replace $(O\colon N)$ by "every ideal of N contained in $(O\colon N)'$", respectively $(O\colon N)''$.
With these alterations we now have

THEOREM 2.11. *Let R be a KA-radical on a universal class $\mathfrak{U}$ of near-rings such that every semisimple near-ring is zerosymmetric. Then*
(i) *R is hypersolvable is equivalent to each of* (ii)$'$ *to* (viii)$'$ *and to each of* (ii)$''$ *to* (viii)$''$.

PROOF. (i)$\Rightarrow$(ii)$'$. Since every semisimple near-ring is zero-symmetric, $N/K\in \mathbf{S}$ implies $N_c\subseteq K$ and so, by 1.8, $K\subseteq (K\colon N)'$. If J is any ideal of N contained in $(K\colon N)'$ then $J^2\subseteq N(K\colon N)'\subseteq K$ and so by 2.2, $J\subseteq K$. Hence $K=(K\colon N)'_N$.

(i)⇒(ii)″. From (ii) and (ii)′, $K=(K\colon N)=(K\colon N)'_N$. Hence $K=(K\colon N)\cap\cap(K\colon N)'_N\subseteq(K\colon N)\cap(K\colon N)'\subseteq(K\colon N)=K$, i.e. $K=(K\colon N)''$.

(i)⇒(v)′. Since $N_c\subseteq R(N)$ and $R(N)\subseteq\varphi_I^{-1}R(N/I)$ we have from 1.8

$$R(N)\subseteq\varphi_I^{-1}R(N/I)\cap(R(N)\colon U).$$

The proof of the reverse inclusion is similar to that in (i)⇒(v).

The proof that (i)⇒(v)″ is similar. All the other implications are proved in a way similar to the corresponding results in 2.3.

The notation introduced in 2.7 also has right and two-sided analogues. Thus for each triple (I, U, T) with $I\lhd N$ and $T, U\subseteq N$ we define

$$X'(I, U, T)=[\varphi_I^{-1}R(N/I)\cap(T\colon U)']_N$$

$$X''(I, U, T)=[\varphi_I^{-1}R(N/I)\cap(T\colon U)'']_N.$$

By methods similar to those used in 2.7, 2.10 and 2.11 we obtain

PROPOSITION 2.13. *The results obtained from those in* 2.7, 2.10 *and* 2.11 *by replacing X by X', respectively X by X'', and* (C1) *by*

$$\text{(C1)}'\quad U\cdot R(N)\subseteq T$$

respectively

$$\text{(C1)}''\quad (R(N)\cdot U)\cup(U\cdot R(N))\subseteq T$$

are valid.

REMARK 2.14. If for a triple (I, U, T) the simplified formula

$$R(N)=\varphi_I^{-1}R(N/I)\cap(T\colon U)$$

holds and if $U\cdot R(N)$ is contained in T then the two-sided quotient formula simplifies to

$$R(N)=\varphi_I^{-1}R(N/I)\cap T\colon U)'',$$

since, in this case,

$$R(N)\subseteq\varphi_I^{-1}R(N/I)\cap(T\colon U)\cap(T\colon U)'$$

$$\subseteq\varphi_I^{-1}R(N/I)\cap(T\colon U)=R(N).$$

The above formulae, $R(N)=X(I, U, T)$, $R(N)=X'(I, U, T)$, and $R(N)=$ $=X''(I, U, T)$ are generalizations of Andrunakievic's characterization formula for hypersolvable ring radicals with hereditary radical class ([2]), (see also [13], Section 11) to the nonhereditary case. There are problems in the practical use of our formulae because of the need to find suitable pairs (T, U) which do not involve $R(N)$. As we shall see in the next section, these problems disappear in the case of hereditary radical classes.

3. Supernilpotency

A radical R on a universal class $\mathfrak{U}$ of near-rings is *supernilpotent* if it is hypersolvable and has a hereditary radical class. Recall that for a KA-radical R on a universal class $\mathfrak{U}$ of near-rings the radical class is hereditary if and only if for every $N\in\mathfrak{U}$ and every $I\lhd N$, $R(N)\cap I\subseteq R(I)$.

THEOREM 3.1. *For a hypersolvable KA-radical R on a universal class $\mathfrak{U}$ of near-rings the following are equivalent:*

(i) *R has a hereditary radical class;*

(ii) *$R(N)\cap I\subseteq(R(I)\colon N)$ for all $N\in\mathfrak{U}$ and $I\lhd N$;*

(iii) *$R(N)\cap I\subseteq(R(I)\colon U)$ for all $N\in\mathfrak{U}$, and $I\lhd N$ and all subsets U of N;*

(iv) *$R(N)\cap I\subseteq(R(I)\colon I)$ for all $N\in\mathfrak{U}$ and $I\lhd N$;*

(v) *$R(N)\cap I=((R(I)\colon I)\cap I)_N=R(I)_N$ for all $N\in\mathfrak{U}$ and $I\lhd N$.*

PROOF. (ii)⇒(iii)⇒(iv) and (v)⇒(iv) are trivial.

(i)⇒(ii). Since R has hereditary radical class, $R(N)\cap I\subset R(I)$. Applying 2.3 (iii) we have

$$R(N)\cap I=(R(N)\colon N)\cap I.$$

But

$$((R(N)\colon N)\cap I)N\subseteq R(N)\cap I\subseteq R(I)$$

and so

$$(R(N)\colon N)\cap I\subseteq(R(I)\colon N).$$

It follows that $R(N)\cap I\subset(R(I)\colon N)$ as required.

(i)⇒(v). Since (i)⇒(iv) we see that $R(N)\cap I\subseteq(R(I)\colon I)$. Let K be an ideal of N contained in both I and $(R(I)\colon I)$. Then $K^2\subseteq(R(I)\colon I)I\subseteq R(I)$ and hence by 2.2, $K\subseteq R(I)$. Since **R** is hereditary, $K\in\mathbf{R}$ and hence $K\subseteq R(N)$. It follows that $R(N)\cap I$ is the largest ideal of N contained in $(R(I)\colon I)\cap I=R(I)$. Hence we have (v). Finally, (iv)⇒(i) because $R(N)\cap I\subseteq(R(I)\colon I)=R(I)$.

Another characterization, generalizing Andrunakievich's formula [2] to left quotients and near-rings, is given by

THEOREM 3.2. *For a KA-radical R on a universal class $\mathfrak{U}$ of near-rings the following are equivalent.*

(i) *R is supernilpotent;*

(ii) *$R(N)=X(I,U,T)$ for all $N\in\mathfrak{U}$ and every triple (I,U,T), with $I\lhd N$, $I\subseteq U\subseteq N$, satisfying $R(N)\cdot U\subseteq T$ and $T\cap I\subseteq R(I)$ (and there is at least one such triple);*

(iii) *$R(N)=X(I,I,R(I)+R(N)_c)=\varphi_I^{-1}R(N/I)\cap(R(I)+R(N)_c\colon I)_N$ for all $N\in\mathfrak{U}$ and $I\lhd N$;*

(iv) *$R(N)=X(I,U,T)$ for all $N\in\mathfrak{U}$ and $I\lhd N$, for suitable subsets T, U of N satisfying $I\subseteq U\subseteq N$ and $T\cap I\subseteq R(I)$.*

PROOF. (i)⇒(ii). By 2.7 it is enough to find a suitable ideal K for which (C3) is satisfied. Now $T\cap I\subseteq R(I)$ implies $((T\colon U)\cap I)_N^2\subseteq(T\colon U)I\cap I\subseteq T\cap I\subseteq R(I)$.

Applying 2.2 we deduce that $((T\colon U)\cap I)_N \subseteq R(I)$ and, since $\mathbf{R}$ is hereditary, $((T\colon U)\cap I)_N \in \mathbf{R}$. It follows that

$$\begin{aligned} I\cap X(I, U, T) &= I\cap(\varphi_I^{-1} R(N/I)\cap(T\colon U))_N \\ &= (I\cap(T\colon U))_N \end{aligned}$$

by 1.9 and so $I\cap X(I, U, T)\subseteq R(N)$.

Since $X(I, U, T)\subseteq \varphi_I^{-1} R(N/I)$ we then have

$$X(I, U, T)/X(I, U, T)\cap I \subseteq R(N/I)$$

and so $X(I, U, T)/X(I, U, T)\cap I \in \mathbf{R}$ because $\mathbf{R}$ is hereditary.

Taking $K=X(I, U, T)\cap I$ we then have 2.7 (iii) and hence $R(N)=X(I, U, T)$ as required.

(That there must be one such triple, at least, is demonstrated by the triple $(I, N, R(N))$).

(i)$\Rightarrow$(iii). From (i), if $I\lhd N$, $R(N)\cap I\subseteq R(I)$ and so using 1.8,

$$\begin{aligned} R(N)\cdot I &\subseteq R(N)\cap(I+R(N)_c) = (R(N)\cap I)+R(N)_c \\ &\subseteq R(I)+R(N)_c. \end{aligned}$$

Now

$$(R(I)+R(N)_c)\cap I = R(I)+(R(N)_c\cap I) \subseteq R(I)$$

and so with $T=R(I)+R(N)_c$, $U=I$ we have $R(N)\cdot U\subseteq T$ and $T\cap I\subseteq R(I)$. Since (i) implies (ii) we have

$$R(N) = X(I, I, R(I)+R(N)_c)$$

as required.

Observe that (ii) implies (iv) and that (iii) implies (iv) are trivial.

(iv)$\Rightarrow$(i). Since (iv) holds we have R hypersolvable by 2.3 (vi). Also using 1.9,

$$R(N)\cap I = (\varphi_I^{-1} R(N/I)\cap(T\colon U))_N\cap I = (I\cap(T\colon U))_N.$$

Since $I\subseteq U\subseteq N$ we obtain from this

$$(R(N)\cap I)I \subseteq ((T\colon U)\cap I)I \subseteq T\cap I \subseteq R(I),$$

i.e. $R(N)\cap I\subseteq(R(I)\colon I)$.

It follows from 3.1 (iv) that R has a hereditary radical class.

COROLLARY 3.3. *If R is a supernilpotent KA-radical on a universal class $\mathfrak{U}$ of near-rings then the semisimple class $\mathbf{S}$ satisfies the following: if $N\in\mathbf{S}$ and $I\lhd N$ with $N_c\subseteq I$ then $\{O\}$ is the only subset of $R(I)$ which is both an ideal in I and a left ideal in N.*

PROOF. If $N\in\mathbf{S}$ then $R(N)=(0)$ so if T is a subset of $R(I)$ which is both an ideal in I and a left ideal in N, we can apply 3.2 (ii) with $U=I$ to get

$$X(I, I, T) = (\varphi_I^{-1} R(N/I)\cap(T\colon I))_N = (0).$$

But by 1.6

$$T\subseteq I\cap(T\colon I) \subseteq \varphi_I^{-1} R(N/I)\cap(T\colon I) = R(N) = (0).$$

Hence $T=(0)$ as required.

When $\mathfrak{U}$ consists only of zero symmetric near-rings then $\mathbf{S}$ is "nearly hereditary" in the sense that this property will hold for all ideals I of each $N \in \mathbf{S}$. This "nearly hereditary" property of $\mathbf{S}$ in the zero symmetric case is stronger than the following which holds for every KA-radical with hereditary radical class:

if $N \in \mathbf{S}$ and $I \triangleleft N$ then (0) is the only ideal of N contained in $R(I)$.

This result is direct consequence of the hereditary nature of $\mathbf{R}$. However, even for zero symmetric near-rings, there is no possibility, generally, of extending this "nearly hereditary" property to a "hereditary" property. To see this we need only consider the nil radical (cf. the comments to 4.4).

The results in this section have demonstrated conditions under which a (hypersolvable) KA-radical will have a hereditary radical class. If we restrict attention to universal classes of zero symmetric near-rings we can obtain similar results starting with more general maps than KA-radicals.

THEOREM 3.4. *Let $\mathfrak{U}$ be a universal class of zero symmetric near-rings and R be a mapping which assigns to each $N \in \mathfrak{U}$ an ideal $R(N)$ of N. Then the following are equivalent:*

(i) *R is a supernilpotent KA-radical on $\mathfrak{U}$;*

(ii) (d1) *$R(N/R(N))=(0)$ for all $N \in \mathfrak{U}$ and*

(d2) *$R(N)=\varphi_I^{-1}R(N/I)\cap(J(I)\colon I)$ for all $N \in \mathfrak{U}$ and all $I \triangleleft N$ where $J(I)$ is the maximal ideal of N contained in $R(I)$;*

(iii) (d1) *and*

(d3) *$R(N)=\varphi_I^{-1}R(N/I)\cap(L(I)\colon I)$ for all $N \in \mathfrak{U}$ and all $I \triangleleft N$ where $L(I)$ is the maximal left ideal of N contained in $R(I)$.*

PROOF. (i)$\Rightarrow$(ii). Since a KA-radical is an H-radical (d1) is immediate. Since the near-rings in $\mathfrak{U}$ are zero-symmetric,

$$R(N)\cdot I \subseteq R(N)\cap I \subseteq J(I).$$

Hence we have (d2) from 3.2 (ii) on taking $U=I$ and $T=J(I)$.

(i)$\Rightarrow$(iii). This is similar to the above using $J(I)\subseteq L(I)$ and $T=L(I)$.

(ii)$\Rightarrow$(i). Since (d2) implies $\varphi_I R(N)\subseteq R(\varphi_I N)$, this together with (d1) shows that R is an H-radical. If $N \in \mathbf{S}$ and I is a radical ideal of N then $I=R(I)$. But then by (d2) we have (since $I=J(I)$)

$$I \subseteq (I\colon I)\cap\varphi_I^{-1}R(N/I) = R(N) = (0).$$

Then if $N \in \mathfrak{U}$ and $I \triangleleft N$ with $I=R(I)$ we have

$$I/R(N)\cap I \in \mathbf{R} \quad\text{and}\quad I/R(N)\cap I \triangleleft N/R(N) \in \mathbf{S}$$

from which $I/(R(N)\cap I)=(0)$. Thus $R(N)$ contains all radical ideals of N. Now suppose that $N \in \mathfrak{U}$ with $N^2=(0)$. By (d2) with $I=N$, $T=J(I)$ we have

$$R(N) = N\cap(T\colon N) = N.$$

Hence $N^2=(0)$ implies $N \in \mathbf{R}$. Now from (d2), $R(N)\subseteq(J(I)\colon I)$ so $R(N)\cdot I\subseteq$ $\subseteq J(I)\subseteq R(I)$. Taking $I=R(N)$ we get $R(N)^2\subseteq R(R(N))$. This means that

$R(N)/R(R(N))$ is a zero near-ring and hence radical. Since it is also semisimple it must be zero and so $R(N)=R(R(N))$. These last two results convert an H-radical into a KA-radical. Since every zero near-ring is radical it follows from 2.1 that R is hypersolvable. By 3.2 (iv) it then follows that **R** is hereditary and we have the result.

(iii)$\Rightarrow$(i) is proven similarly.

We now turn to methods of computing $R(N)$ for a KA-radical with hereditary radical class from a knowledge of the radicals of near-rings related to N. Our result is valid irrespective of whether the radical is hypersolvable.

THEOREM 3.5. *If R is a KA-radical with hereditary radical class on a universal class $\mathfrak{U}$ of near-rings then for every $N\in\mathfrak{U}$ and every ideal I of N, $R(N)$ is the largest ideal P of N contained in $\varphi_I^{-1}R(N/I)$ which satisfies $P\cap I\subseteq R(I)$.*

PROOF. Certainly $R(N)$ is an ideal of N contained in $\varphi_I^{-1}R(N/I)$ and since the radical is hereditary, $R(N)\cap I\subseteq R(I)$. Now let K be an ideal of N contained in $\varphi_I^{-1}R(N/I)$ and satisfying $K\cap I\subseteq R(I)$. Since **R** is hereditary, $K\cap I\in\mathbf{R}$ and so $K\cap I\subseteq R(N)$. Then

$$K/K\cap R(N)\simeq\frac{K/K\cap I}{K\cap R(N)/K\cap I}.$$

Now $K/K\cap I\cong(K+I)/I\in\mathbf{R}$ since $(K+I)/I$ is an ideal of N/I contained in $R(N/I)$. Because **R** is hereditary, $K\cap R(N)\in\mathbf{R}$. Hence by the extension property, 1.1 (iii), we have $K\in\mathbf{R}$ and hence $K\subseteq R(N)$.

We now turn again to a consideration of right and two-sided quotients. The statements (ii) to (v) of 3.1 can be modified to yield statements (ii)′ to (v)′, respectively (ii)″ to (v)″, in the following way:

(ii) — replace $(R(I)\colon N)$ by $(R(I)+N_c\colon N)'$, respectively by $(R(I)+N_c\colon N)''$;

(iii) — similarly;

(iv), (v) — replace $(R(I)\colon I)$ by $(R(I)\colon I)'$, respectively by $(R(I)\colon I)''$.

With these modifications and assuming $N_c\subseteq R(N)$, we have the following which is proved in a similar way to 3.1 observing that it follows that both

$$N((R(N)\colon N)'\cap I)\subseteq R(N)\cap(I+N_c)=(R(N)\cap I)+N_c$$

and

$$(R(I)+N_c\colon I)'=(R(I)+I_c\colon I)'=(R(I)\colon I)'$$

are true.

THEOREM 3.6. *If R is a hypersolvable radical on a universal class $\mathfrak{U}$ of near-rings such that every semisimple near-ring is zero-symmetric then*

(i) *R has a hereditary radical class is equivalent to each of* (ii)′ *to* (v)′ *and to each of* (ii)″ *to* (v)″.

Modifications to (ii) to (iv) of 3.2 to yield (ii)′ to (iv)′ and (ii)″ to (iv)″, respectively, can be made as follows:

(ii) Replace X by X' and $R(N)U\subseteq T$ by $UR(N)\subseteq T$, respectively X by X'' and $R(N)U\subseteq T$ by $R(N)U\cup U\cdot R(N)\subseteq T$;

(iii) replace $X(I, I, R(I)+N_c)$ by $X'(I, I, R(I))$, respectively $X''(I, I, R(I)+N_c)$;

(iv) replace X by X', respectively X''.

Then

THEOREM 3.7. *For a KA-radical R on a universal class $\mathfrak{U}$ of near-rings for which every semisimple near-ring is zero-symmetric,*
(i) *R is supernilpotent*
is equivalent to each of (ii)′ *to* (iv)′ *and to each of* (ii)″ *to* (iv)″.

PROOF. The equivalence of (i) to each of (ii)′ to (iv)′ is proved similarly to 3.2, observing that with our assumptions on R, $N_c \subseteq R(N)$ and

$$IR(N) \subseteq I \cap (R(N)+I_c) = (I \cap R(N)) + I_c \subseteq R(I)+I_c = R(I).$$

The equivalence of (i) to each of (ii)″ to (iv)″ follows by intersecting left and right quotients. For (iii)″ we observe that

$$\begin{aligned} R(N) &= \varphi_I^{-1} R(N/I) \cap (R(I)+N_c\colon I)_N \\ &\qquad \cap (\varphi_I^{-1} R(N/I) \cap)(R(I)\colon I)'_N \\ &\subseteq (\varphi_I^{-1} R(N/I) \cap (R(I)+N_c\colon I)'')_N \\ &\subseteq (\varphi_I^{-1} R(N/I) \cap (R(I)+N_c\colon I)_N = R(N). \end{aligned}$$

The result (iv)″⇒(i) has the same proof as (iv)⇒(i).

Finally 3.4 has its right and two-sided quotient version proved in an essentially similar way.

THEOREM 3.8. *If $\mathfrak{U}$ is a universal class of zero-symmetric near-rings, the right and two-sided analogues of* 3.4 *obtained by replacing $(J(I)\colon I)$, respectively $(L(I)\colon I)$ by $(J(I)\colon I)'_N$ or $(J(I)\colon I)''_N$, respectively $(L(I)\colon I)'_N$ or $(L(I)\colon I)''_N$ hold.*

4. Weakly special classes

For associative rings Rjabuhin [12] has defined weakly special classes **K** by the following properties.

(1) $N \in \mathbf{K} \Rightarrow N$ contains no nilpotent ideal (i.e. N is semiprime);

(2) **K** is hereditary;

(3) $I \lhd N, I \in \mathbf{K}, (0\colon I)'' = 0 \Rightarrow N \in \mathbf{K}$.

An obvious example of a weakly special class is the semisimple class corresponding to a supernilpotent radical.

If we want to generalize these classes to near-rings we will clearly need (1). Since the semisimple classes of near-ring radicals are not generally hereditary we will need to modify (2) for the near-ring case. The comments after 3.3 suggest one type of modification required. There remains the need to consider analogues of (3) for the near-ring case.

PROPOSITION 4.1. *Let N be a near-ring and I, J be ideals of N with $I \cap J = (0)$. If P is any ideal of I containing no nonzero ideals of N then $\varphi_J P$ is an ideal of $\varphi_J I$ containing no non-zero ideals of $\varphi_J N$.*

PROOF. Let K be an ideal of $\varphi_J N = N/J$ contained in $\varphi_J P$ and $n \in (\varphi_J^{-1} K) \cap I$. Then $\varphi_J(n) \in K \subseteq \varphi_J P$. Hence for some $p \in P$, $\varphi_J(n) = \varphi_J(p)$ and so $n - p \in J \cap I = (0)$. It follows that $n \in P$. Then $(\varphi_J^{-1} K) \cap I$ is an ideal of N contained in P and thus

$(\varphi_J^{-1}K)\cap I=(0)$. Since $P\subseteq I$, $\varphi_J P\subseteq\varphi_J I$ and so $k\in K$ implies $k=\varphi_J(x)$ for some $x\in I$. Then $x\in(\varphi_J^{-1}K)\cap I=(0)$. Hence $K=0$ as required.

PROPOSITION 4.2. *Let R be a supernilpotent KA-radical on a universal class $\mathfrak{U}$ of near-rings. Suppose also that R satisfies*

(*) *$R(N)=0\Rightarrow R(\varphi N)$ is zero-symmetric for every homomorphism φ defined on N.*

Then if N is semisimple, $N/(O\colon I+N_c)$ is semisimple for every ideal I of N.

PROOF. Let N be semisimple and I be an ideal of N. Then $(I\cap(O\colon I+N_c))$ is an ideal of N (by 1.6) which is nilpotent and hence 0. From the remarks following 3.3, the only ideal of N contained in $R(I)$ is the zero ideal and so applying 4.1 with $P=R(I)$, $J=(O\colon I+N_c)$ we see that 0 is the only ideal of $\varphi_J N=N/(O\colon I+N_c)$ which is contained in $\varphi_J R(I)$. But φ_J is an isomorphism on I and so $\varphi_J R(I)=R(\varphi_J I)$ and hence since $R(\varphi_J N)\cap\varphi_J I$ is an ideal of $\varphi_J N$ contained in $R(\varphi_J I)$ it must be the zero ideal. By 3.2 (iii) we have

$$\begin{aligned} R(\varphi_J N) &\subseteq \big((R(\varphi_J N)\cap\varphi_J I)+(R(\varphi_J N))_c\colon \varphi_J I+(\varphi_J N)_c\big) \\ &= \big(0\colon \varphi_J I+(\varphi_J N)_c\big) \end{aligned}$$

by our assumption that $(R(\varphi_J N))_c=0$. Since $\varphi_J(N_c)=(\varphi_J N)_c$ we obtain

$$\begin{aligned} \varphi_J^{-1}\big(0\colon \varphi_J I+(\varphi_J N)_c\big) &= \big\{n\in N\,|\,\varphi_J n(\varphi_J i+\varphi_J c)=0 \text{ for all } i\in I, c\in N_c\big\} \\ &= \{n\in N\,|\,n(i+c)\in J,\ i\in I,\ c\in N_c\} \\ &= (J\colon I+N_c) \\ &= \big((0\colon I+N_c)\colon I+N_c\big). \end{aligned}$$

Thus

$$R(\varphi_J N)\subseteq\frac{\big((0\colon I+N_c)\colon I+N_c\big)}{(0\colon I+N_c)}.$$

We need only prove this factor ideal to be zero to obtain the result.

Let $n\in N$, $i\in I$, $c\in N_c$. Then

$$n(i+c)=n(i+c)-nc+nc\in I+N_c$$

and $n\in((O\colon I+N_c)\colon I+N_c)$ if and only if $n(I+N_c)$ is contained in $(O\colon I+N_c)\cap(I+N_c)=(O\colon I+N_c)\cap I=0$, i.e., if and only if $n\in(O\colon I+N_c)$. It follows that $R(\varphi_J N)=0$.

The condition (*) in this last result is satisfied in many important cases, for example in any universal class of zero-symmetric near-rings, for those radicals whose semisimple near-rings are zero-symmetric (in which case N and φN are zero-symmetric) or for those radicals whose radical near-rings are zero-symmetric.

THEOREM 4.3. *Let R be a supernilpotent KA-radical on a universal class $\mathfrak{U}$ of near-rings and suppose R satisfies (*). If I is an ideal of $N\in\mathfrak{U}$ such that $R(I)$ contains no non-zero ideal of N then $N/(R(N)_c\colon I+N_c)$ is semisimple.*

PROOF. From 3.2 and our assumption on $R(I)$ we have $R(N)\subseteq ((R(N)\cap I)+R(N)_c\colon I+N_c)=(R(N)_c\colon I+N_c)$. Consider the subset

$$(0\colon (R(N)+I+N_c)/R(N)) \quad \text{of} \quad N/R(N).$$

By 1.6

$$(0\colon (R(N)+I+N_c)/R(N)) = \left(0\colon \frac{R(N)+I}{R(N)} + \left(\frac{N}{R(N)}\right)_c\right)$$

is an ideal of $N/R(N)$ and so its preimage J is an ideal of N. Then $R(N)\subseteq J$ and

$$J = \{n\in N\colon\ n(r+i+c)\in R(N), \quad \text{for all} \quad r\in R(N),\ i\in I,\ c\in N_c\}.$$

Now $n(r+i+c)=n(r+i+c)-n(i+c)+n(i+c)$ is in $R(N)+n(I+N_c)$ and so $J=(R(N)\colon I+N_c)$. Also, $n(i+c)-nc+nc\in I+N_c$ and hence $J=(R(N)\cap(I+N_c)\colon I+N_c)$. If $r\in R(N)\cap(I+N_c)$ then $r=i+c$ for some $i\in I$, $c\in N_c$ and so $rO=iO+c$. It follows that $r-rO=i-iO\in I$ and so $r\in I+R(N)_c$ from which $R(N)\cap(I+N_c)= =R(N)\cap I+R(N)_c$. Our assumption on $R(I)$ coupled with the hereditary radical class implies $R(N)\cap I=O$ and so $J=(R(N)_c\colon I+N_c)$ and $(R(N)_c\colon I+N_c)$ is an ideal of N containing $R(N)$. Then

$$N/(R(N)_c\colon I+N_c) \simeq \frac{N/R(N)}{(R(N)_c\colon I+N_c)/R(N)} \simeq \frac{N/R(N)}{\left(0\colon \frac{R(N)+I}{R(N)} + \left(\frac{N}{R(N)}\right)_c\right)}$$

which is semisimple by 4.2.

We have a condition in this theorem which, on the assumption of (*) can be used to replace (3) of the definition of weakly special classes.

DEFINITION 4.4. A class $\mathbf{K}$ of near-rings from a universal class $\mathfrak{U}$ is *weakly special* in $\mathfrak{U}$ if it satisfies

(1) $N\in\mathbf{K}\Rightarrow N$ is semiprime;

(2) $N\in\mathbf{K}$, $I\lhd N\Rightarrow I\mathbf{K}$ contains no non-zero ideal of N;

(3) $N\in\mathfrak{U}$, $I\lhd N$ such that $I\mathbf{K}$ contains no non-zero ideal of $N\Rightarrow$ $\Rightarrow N/(N\mathbf{K}_c\colon I+N_c)\in\mathbf{K}$.

(For a near-ring N, $N\mathbf{K}$ denotes the intersection of all ideals I of N with $N/I\in\mathbf{K}$.)

The comments following 3.3 and Theorem 4.3 above together imply that the semi-simple classes of supernilpotent KA-radicals which satisfy (*) are weakly special. For associative rings, it is known that $K\lhd I\lhd N$ and I/K semiprime imply that K is an ideal in N (see eg. [14, prop. 23]). Application of this result to $K=I\mathbf{K}$ yields that for associative rings (1) and (2) are equivalent to (1) and the following weak form of heredity of $\mathbf{K}$:

(2*) $\qquad N\in\mathbf{K},\ I\lhd N$ imply $I\in\overline{\mathbf{K}}$

(where $\overline{\mathbf{K}}$ denotes the subdirect closure of $\mathbf{K}$ in $\mathfrak{U}$). For near-rings (even in the zero-symmetric case) such an equivalence does not hold: it can be verified that (2*) implies the heredity of $\bar{\mathbf{K}}$; but the upper nil radical (which is a supernilpotent KA-radical) has a nonhereditary semisimple class (cf. [6, example 5.4] — since the example is finite, the upper nil radical coincides there with J).

If we consider zero-symmetric near-rings, then condition (3) of 4.4 is equivalent to a near-ring analogue of Rjabuhin's definition:

PROPOSITION 4.5. *If* $N\mathbf{K}_c=0$ *for all* $N\in\mathfrak{U}$ *then* 4.4 (3) *can be replaced with* (3)′ $N\in\mathfrak{U}$, $I\lhd N$, $I\mathbf{K}$ *contains no non-zero ideal of* N,

$$(0\colon I+N_c) = 0 \Rightarrow N\in\mathbf{K}.$$

PROOF. Clearly (3)$\Rightarrow$(3)′ when $N\mathbf{K}_c=0$. Assume (3)′ and let I be an ideal of $N\in\mathfrak{U}$ for which $I\mathbf{K}$ contains no non-zero ideal of N. Since $I\cap(O\colon I+N_c)$ is a nilpotent ideal of N contained in I it follows that $I\cap(O\colon I+N_c)=0$ and so

$$(I+(0\colon I+N_c))/(0\colon I+N_c) = \bar{I}$$

is isomorphic to I and then by 4.1, $\bar{I}$ contains no non-zero ideal of $\bar{N}=N/(O\colon I+N_c)$. We will obtain (3) by showing that $(O\colon \bar{I}+\bar{N}_c)=O$ and applying (3)′. The preimage J of $(O\colon \bar{I}+\bar{N}_c)$ in N is

$$\begin{aligned} J &= \{n\in N\colon \bar{n}(\bar{I}+N_c) = 0\} \\ &= \{n\in N\colon n(I+N_c+(0\colon I+N_c)) \subseteq (0\colon I+N_c)\} \\ &= \{n\in N\colon n(I_0+(0\colon I+N_c)+N_c) \subseteq (0\colon I+N_c)\} \end{aligned}$$

recalling that I_O is the zero symmetric part of I. If $i_O\in I_O$, $a\in(O\colon I+N_c)$ and $c\in N_c$ then

$$n(i_o+a+c) = n(i_o+a+c)-n(a+c)+n(a+c)-nc+nc\in I_o+(0\colon I+N_c)+N_c$$

and so $n(i_o+a+c)\in(O\colon I+N_c)$ if and only if

$$n(i_o+a+c)-n(a+c)+nc\in(0\colon I+N_c)\cap(I_o+N_c) = 0.$$

This means that $nc=n(a+c)-n(i_O+a+c)\in I_O$, i.e. $nc=0$. In particular, when $a=0$ we have

$$\begin{aligned} n(i_o+c)\in(0\colon I+N_c) \Rightarrow n(i_o+c) &= n(i_o+c)-nc+nc \\ &= n(i_o+c)-nc \\ &\in I_o\cap(o\colon I+N_c) = O. \end{aligned}$$

Hence $J\subseteq(O\colon I+N_c)$ and so $(O\colon \bar{I}+\bar{N}_c)=O$.

Now we turn to the use of weakly special classes in the definition of a radical.

PROPOSITION 4.6. *Let* $\mathfrak{U}$ *be a universal class of near-rings and* $\mathbf{K}$ *be a subclass of* $\mathfrak{U}$ *satisfying Condition* 4.4 (2). *Then the class* $\mathfrak{R}\mathbf{K}$ *of all near-rings* N *in* $\mathfrak{U}$ *having no non-zero homomorphic image in* $\mathbf{K}$ *is a KA-radical class (called the upper radical class determined by* $\mathbf{K}$ *in* $\mathfrak{U}$*).*

PROOF. 4.4 (2) implies that $\mathbf{K}$ is a regular class, i.e. satisfies $N\in\mathbf{K}$, $I\lhd N\Rightarrow I\mathbf{K}\neq I$.

It follows that the corresponding upper radical class $\mathfrak{R}\mathbf{K}$ is a KA-radical class (see for example [5] or [14] the result is formulated there for rings, but the proof holds for Ω-groups).

Finally, we will consider the bijective correspondence between weakly special classes and supernilpotent KA-radicals in the two (extreme) cases of zero-symmetric semisimple near-rings and zero-symmetric radical near-rings.

THEOREM 4.7. *Let $\mathfrak{U}$ be a universal class of near-rings. Then the following holds:*

(A) $\mathbf{R}$ *is the radical class of a supernilpotent KA-radical R on $\mathfrak{U}$ whose semisimple class consists only of zero-symmetric near-rings if and only if* $\mathbf{R}$ *is the upper radical class $\mathfrak{R}\mathbf{K}$ determined by some weakly special class $\mathbf{K}\subseteq\mathfrak{U}$ consisting of zero-symmetric near-rings.*

(B) $\mathbf{R}$ *is the radical class of a supernilpotent KA-radical on $\mathfrak{U}$ whose radical class consists only of zero-symmetric near-rings if and only if* $\mathbf{R}$ *is the upper radical class $\mathfrak{R}\mathbf{K}$ determined by some weakly special class $\mathbf{K}\subseteq\mathfrak{U}$ whose subdirect closure contains all near-rings $N\in\mathfrak{U}$ which satisfy $(0\colon N_c)=0$.*

In both cases, the corresponding radical is given by $R(N)=N\mathbf{K}$ on every $N\in\mathfrak{U}$ and the corresponding semisimple class is the subdirect closure of $\mathbf{K}$ *in* $\mathfrak{U}$.

PROOF. If R is a KA-radical on $\mathfrak{U}$ which satisfies the conditions of (A) resp (B), then R satisfies the conditions of Theorem 4.3 and the corresponding semisimple class $\mathbf{S}$ is weakly special. Hence the theorem holds for $\mathbf{K}=\mathbf{S}$.

If conversely $\mathbf{R}$ is of the form $\mathbf{R}=\mathfrak{R}\mathbf{K}$ for some weakly special class $\mathbf{K}$ in $\mathfrak{U}$, $\mathbf{R}$ is the radical class of a KA-radical R on $\mathfrak{U}$ by Proposition 4.6, which is hypersolvable by condition (1) for weakly special classes. Let us first compute $R(N)$. If I is an ideal of $N\in\mathfrak{U}$ with $N/I\in\mathbf{K}$, then $(R(N)+I)/I$ is a radical ideal of N/I for which therefore $((R(N)+I/I)\mathbf{K}=(R(N)+I)/I$ is an ideal of N/I. Condition (2) in Definition 4.4 now implies $(R(N)+I)/I=0$, i.e., $R(N)\subseteq I$.

To obtain $R(N)=N\mathbf{K}$ it suffices to show that $N\mathbf{K}$ belongs to $\mathbf{R}$ for every $N\in\mathfrak{U}$. Thus let J be an ideal of $N\mathbf{K}$ sending $N\mathbf{K}$ into $\mathbf{K}$ and denote by J' the largest ideal of N contained in J. Then we have

$$\frac{N\mathbf{K}/J'}{J/J'}\simeq\frac{N\mathbf{K}}{J}\in\mathbf{K}$$

implying $(N\mathbf{K}/J')\subseteq J/J'$.

By construction J/J' contains no non-zero ideal of N/J' so that (3) of Definition 4.4 can be applied to the ideal $N\mathbf{K}/J'$ of N/J' yielding

$$((N/J')K_c\colon\ NK/J'+N_c/J')\in\mathbf{K}.$$

Denoting by X the set of all elements $c\in N_c$ with $c+J'\in(N/J)\mathbf{K}_c$, a straightforward calculation shows that the preimage of $((N/J')\mathbf{K}_c\colon N\mathbf{K}/J'+N_c/J')$ in N is $(X+J'\colon N\mathbf{K}+N_c)$. This implies that $(X+J'\colon N\mathbf{K}+N_c)$ is an ideal in N and the isomorphism theorem yields

$$\frac{N}{(X+J'\colon\ N\mathbf{K}+N_c)}\simeq\frac{N/J'}{((N/J')\mathbf{K}_c\colon\ N\mathbf{K}/J'+N_c/J')}\in\mathbf{K}.$$

Thus $N\mathbf{K}$ is contained in $(X+J'\colon\ N\mathbf{K}+N_c)$ and $N\mathbf{K}^2\subseteq X+J'\subseteq X+J$. In case (A) every near-ring in $\mathbf{K}$ is zero-symmetric so that $N\mathbf{K}$ contains N_c for every N in $\mathfrak{U}$. This implies in particular $N_c \subseteq J \subseteq N\mathbf{K}$, i.e. $N_c \subseteq N\mathbf{K}^2 \subseteq X+J=J$ and $(N\mathbf{K}/J)^2=0$. For case (B) consider for an arbitrary $N\in\mathfrak{U}$ the factor-near-ring $\bar{N}=N/(O:\, N_c)$; a straightforward calculation shows that $(\bar{O}:\, \bar{N}_c)=\bar{O}$, i.e. that $\bar{N}\mathbf{K}=\bar{O}$, implying $N\mathbf{K}\subseteq(O:\, N_c)$. Hence in case (B), $N\mathbf{K}$ is always zero-symmetric and we obtain $N\mathbf{K}^2\subseteq J$, i.e. again $(N\mathbf{K}/J)^2=0$. Thus in both cases (A) and (B), $N\mathbf{K}/J$ is a nilpotent near-ring belonging to $\mathbf{K}$. This, however, implies $N\mathbf{K}=J$ and we have shown that $N\mathbf{K}$ belongs to $R\mathbf{K}$.

The property $R(N)=N\mathbf{K}$, $\forall N\in\mathfrak{U}$ which we have shown clearly implies that the semisimple class of R is the subdirect closure of $\mathbf{K}$ in $\mathfrak{U}$.

From this fact we easily deduce that in case (A) the semisimple class consists of zero-symmetric near-rings. For case (B) we have shown above that $R(N)=N\mathbf{K}$ is always zero-symmetric, i.e. $\mathbf{R}=\mathfrak{R}\mathbf{K}$ contains only zero-symmetric near-rings. It only remains to show the heredity of $\mathbf{R}$ in each case. To this end let us consider a near-ring N in $\mathfrak{R}\mathbf{K}$ and an ideal I of N. As above for $N\mathbf{K}$, let J be an ideal of I sending I into $\mathbf{K}$ and J' the largest ideal of N contained in J. As above we obtain:

$$\frac{N/J'}{((N/J')\mathbf{K}_c:\ I/J'+N_c/J')}\in\mathbf{K}.$$

Because N belongs to $\mathfrak{R}\mathbf{K}$, this implies

$$N/J' = (N_c/J':\ I/J'+N_c/J')$$

and thus $(I/J')^2$ is contained in I_c/J'.

In case (A) we have $I_c\subseteq I\mathbf{K}\subseteq J$, hence $I^2\subseteq J$.

In case (B) N is zero-symmetric implying $I_c=0$ and $I^2\subseteq J'\subseteq J$. In both cases I/J is a nilpotent near-ring belonging to $\mathbf{K}$ and we obtain $I=J$, i.e. $I\in\mathfrak{R}\mathbf{K}$.

REMARK 4.8. A second type of modification of Rjabuhin's definition for associative rings is suggested by [14, Theorem 43]: condition (2) in Definition 4.4. implies

(2′) $\mathbf{K}$ is regular

(see Prop. 4.6); a straightforward checking shows that Theorem 4.7 and its proof remain valid if we replace in Definition 4.4. Condition (2) by (2′). A further modification which does not affect Theorem 4.7 and its proof may be obtained by replacing (3) in Definition 4.4 by the weaker condition

(3′) $N\in\mathfrak{U}$, $I\lhd N$ such that $I\mathbf{K}$ contains no nonzero ideal of N

$$\Rightarrow N/(N\mathbf{K}_c:\ I+N_c)\in\bar{\mathbf{K}}$$

where $\bar{\mathbf{K}}$ denotes the subdirect closure of $\mathbf{K}$ in $\mathfrak{U}$.

We shall call classes satisfying (1), (2) and (3′) *generalized weakly special classes.*

PROPOSITION 4.9. *Let* $\mathbf{K}\subseteq\mathfrak{U}$ *fulfil* (1) *and* (3) *resp.* (1) *and* (3′) *and suppose either that* $\mathbf{K}$ *contains only zero-symmetric near-rings or that* $\bar{\mathbf{K}}$ *contains all near-rings* N *which satisfy* $(O:\, N_c)=0$. *Then* $\mathbf{K}$ *fulfils* (2) *if and only if it fulfils* (2′).

PROOF. For the implication (2)⇒(2′) see Prop. 4.6. Conversely, under our assumptions on $\mathbf{K}$, if (2′) holds, $\mathfrak{R}\mathbf{K}$ is the radical class of a supernilpotent KA-radical whose semisimple class is $\bar{\mathbf{K}}$ (see 4.8). Hence $\bar{\mathbf{K}}$ fulfils (2) and we obtain (2) for $\mathbf{K}$ using $I\bar{\mathbf{K}}=I\mathbf{K}$ for all $I\lhd N\in\mathfrak{U}$.

Corollary 4.10. *Using in Theorem 4.7 weakly special resp. generalized weakly special classes (both defined either using* (2) *or* (2′)*), one obtains altogether four characterizations of supernilpotent radicals.*

The concept of generalized weakly special classes may be more useful than the weakly special classes themselves, since there are a lot of examples of such classes.

Example 4.11. For zero-symmetric near-rings, Kaarli has introduced special classes (see [7]):

For a homomorphically closed class $\mathfrak{U}$ of zero-symmetric near-rings, a class $\mathbf{K}$ of zero-symmetric near-rings is called $\mathfrak{U}$-special if it satisfies:

(a) all near-rings in $\mathbf{K}$ are prime,

$$(\text{i.e. } A, B \lhd N, A \cdot B = 0 \Rightarrow A = 0 \text{ or } B = 0);$$

(b) $N \in \mathfrak{U} \cap \mathbf{K},\ I \lhd N \Rightarrow I \in \mathbf{K}$;

(c) $K \lhd I \lhd N \in \mathfrak{U},\ I/K \in \mathbf{K} \Rightarrow K \lhd N$ and $N/(K\colon I) \in \mathbf{K}$.

Taking for $\mathfrak{U}$ *the class of all zero-symmetric near-rings, we see that every* $\mathfrak{U}$*-special class is generalized weakly special* ((1) and (2) are direct consequences of (a) and (b), (c) implies $I\mathbf{K}=0$ yielding $N/(O\colon I)=N/(K\colon I)\in \overline{\mathbf{K}}$—here K runs over all ideals of I with $I/K \in \mathbf{K}$).

In order to obtain another class of useful examples, let us recall that many near-ring radicals can be defined by the use of annihilators of N-groups: if to every near-ring N in $\mathfrak{U}$ we have assigned a (possibly empty) class $K(N)$ of N-groups G (with $NG \neq 0$), we obtain a general class $\mathfrak{K}$ on $\mathfrak{U}$, if under

$$(n+I) \cdot g = n \cdot g \qquad \forall n \in N, \quad \forall g \in G,$$

the following holds:

(d) $I \lhd N,\ G \in K(N/I) \Rightarrow G \in K(N)$;

(e) $I \lhd N,\ G \in K(N),\ I \subseteq Ann_N G \Rightarrow G \in K(N/I)$.

It has been shown in [9] (see also [10]) that every general class $\mathfrak{K}$ on $\mathfrak{U}$ defines an H-radical $R_{\mathfrak{K}}$ on $\mathfrak{U}$ by

$$R_{\mathfrak{K}}(N) = \bigcap_{G \in K(N)} Ann_N G \quad \text{for all} \quad N \in \mathfrak{U}.$$

Notice that if every near-ring in $\mathfrak{U}$ can be embedded into a near-ring with right-identity belonging to $\mathfrak{U}$, then every H-radical can be obtained in this way.

A near-ring is $R_{\mathfrak{K}}$-semisimple if and only if it is a subdirect product of $\mathfrak{K}$-primitive near-rings (i.e. near-rings N with a faithful N-group G in $K(N)$). Using this property and applying Theorem 4.3, we obtain:

Example 4.12. *Let* $\mathfrak{K}$ *be a general class on* $\mathfrak{U}$*; then, if the radical* $R_{\mathfrak{K}}$ *is a supernilpotent KA-radical and satisfies* (*), *the class* $\mathbf{K}$ *of all* $\mathfrak{K}$*-primitive near-rings is a generalized weakly special class.*

In the case of zero-symmetric near-rings, we can use Corollary 4.10 to obtain a stronger result:

A radical $R_{\mathfrak{R}}$ on a universal class $\mathfrak{U}$ of zero-symmetric near-rings is supernilpotent if and only if the class of all $\mathfrak{R}$-primitive near-rings is a generalized weakly special class.

Notice that in general we cannot expect the classes of $\mathfrak{R}$-primitive near-rings to be weakly special; for example, the class of all 2-primitive near-rings is not: the ring $\mathbf{Z}$ of the integers satisfies $J_2(I)=I\mathbf{K}=0$ and $(O\colon I)=O$ for every $I\lhd \mathbf{Z}$; but $\mathbf{Z}$ is not 2-primitive.

ACKNOWLEDGEMENT. The first author wishes to acknowledge with thanks the hospitality and financial support provided by Teesside Polytechnic during his visit in 1983 when much of the work contained in this paper was done.

REFERENCES

[1] ANDERSON, T., KAARLI, K. and WIEGANDT, R., On left-strong radicals of near-rings, *Proc Edinburgh Math. Soc.* **31** (1988), 447—456.

[2] ANDRUNAKIEVIČ, V. A., On a characteristic of a hyper-nilpotent radical, *Uspehi Mat. Nauk* **16** (1961), no. 1, 127—130 (in Russian). *MR* **23** # A2442.

[3] BETSCH, G. and KAARLI, K., Supernilpotent radicals and hereditariness of semisimple classes of nearrings, Proc. Conf. on Radical Theory, Eger, 1982; Colloquia Mathematica Soc. János Bolyai 38, 47—58, 1985.

[4] BETSCH, G. and WIEGANDT, R., Nonhereditary semisimple classes of near-rings, *Studia Sci. Math. Hungar.* **17** (1982), 69—75. *MR* **85m**: 16020.

[5] DIVINSKY, N. J., *Rings and radicals,* Mathematical expositions, no. 14, University of Toronto Press, Toronto, Ont., 1965. *MR* **33** # 5654.

[6] KAARLI, K., Classification of irreducible modules over semiprimary near-rings, *Tartu Riikl. Ül. Toimetised,* no. 556 (1981), 47—63 (in Russian). *MR* **82k**: 16047.

[7] KAARLI, K., Special radicals of near-rings, *Tartu Riikl. Ül. Toimetised,* no. 610 (1982), 53—68. *MR* **85c**: 16053.

[8] MLITZ, R., Radicals and semisimple classes of Ω-groups, *Proc. Edinburgh Math. Soc.* (2) **23** (1980), 37—41. *MR* **82e**: 17005.

[9] MLITZ, R., Modules and radicals of universal algebras, *Izv. Vysš. Učebn. Zaved. Matematika,* **1977**, No. 6, 77—85 (Russian). *MR* **58** # 10671. Translation: *Soviet. Math. Izv. VUZ* **21** (1977), 61—67.

[10] MLITZ, R., Radicals and interpolation in universal algebra, Proc. Conf. on Radical Theory, Eger, 1982; Colloquia Math. Soc. János Bolyai 38, 297—331, 1985.

[11] PILZ, G., *Near-rings. The theory and its applications,* North-Holland Mathematics Studies, no. 23, North-Holland Publ. Co., Amsterdam, 1977. *MR* **57** # 9761. Rev Edition: North-Holland Publ. Co., Amsterdam—New York—Oxford, 1983.

[12] RJABUHIN, JU. M., On supernilpotent and special radicals (Russian), Issled Alg. i Mat. Anal, 65—72, Akad. Nauk Moldav. SSR, Kishinev, 1965.

[13] SZÁSZ, F. A., *Radicals of rings,* John Wiley, Chichester—New York, 1981. *MR* **84a**: 16012.

[14] WIEGANDT, R., Radical theory of rings, *Math. Student* **51** (1983), 145—185.

(Received August 6, 1986)

INSTITUT FÜR ANGEWANDTE UND
NUMERISCHE MATHEMATIK
TECHNISCHE UNIVERSITÄT
WIEDNER HAUPTSTRASSE 6—10
A—1040 WIEN
AUSTRIA

DEPARTMENT OF MATHEMATICS
AND STATISTICS
TEESSIDE POLYTECHNIC
MIDDLESBOROUGH
CLEVELAND
TS1 3BA
UNITED KINGDOM

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 259—261

ON A THEOREM OF N. K. BARI

S. A. AVDONIN, S. A. IVANOV and I. JOÓ

The moment problem is a widely investigated theory*; we shall refer only to some papers on this topic. First we mention the work [2] of N. K. Bari, who used moment theorems to characterize the Riesz bases in abstract Hilbert spaces. Since the Riesz basis property allows unconditionally convergent expansions, the Fourier method of mathematical physics is closely related to moment problems. The control problem of a string have thus been investigated via moment theorems (cf. [1], [3], [4], [5][6]). Recently, the first two authors published in [6] some criteria for the controllability of the nets of strings. The paper [6] contains the following theorem without proof. The third author got this result independently.

THEOREM. *Let H be a Hilbert space, $U=(u_n)_{n=1}^{\infty}\subset H$ be a system of normed elements: $\|u_n\|=1$. Define further the linear mapping $B=B_U\colon H\to l_2$, $Bf:=(\langle f, u_n\rangle)_{n=1}^{\infty}$ with domain $D=D_U:=\{f\in H\colon (\langle f, u_n\rangle)_{n=1}^{\infty}\in l_2\}$. Then*

a) *If* $\operatorname{Im} B=l_2$ *then the system U is uniformly minimal. If U is an L-basis in H, i.e. a Riesz basis in the closed linear hull $V(U)$ of U, then* $\operatorname{Im} B=l_2$.

b) *If* $\overline{\operatorname{Im} B}=l_2$ *(the line means the closure in l_2), then U is ω-linearly independent, i.e. $\sum c_n u_n=0$, $(c_n)\in l_2$ implies $c_n=0$ for every n. If U is minimal, then* $\overline{\operatorname{Im} B}=l_2$.

c) *If there exists $U_0\subset U$ such that U_0 is an L-basis and $U\setminus U_0\subset V(U_0)$, then* $\operatorname{Im} B$ *is closed and* $\operatorname{codim} B=\operatorname{card}(U\setminus U_0)$.

PROOF. B is obviously closed. We can assume that U is complete in H, otherwise we can eliminate the component orthogonal to $V(U)$. In this case B is injective.

a) If U is Riesz basis then B is an isomorphism between H and l_2. Conversely, if $\operatorname{Im} B=l_2$ then define $u^n:=B^{-1}\varphi_n$, where (φ_n) is the standard orthonormal basis in l_2. Now

$$(\langle u^n, u_k\rangle)_{k=1}^{\infty} = Bu^n = \varphi_n,$$

hence (u^n) is biorthogonal to (u_n). Since B^{-1} is closed, it is continuous by the closed graph theorem, so

$$\sup_n \|u^n\| \leqq \|B^{-1}\| < \infty$$

and it was to be proved.

---

* The notions and well-known facts used in this paper are given e.g. in the book [7].

1980 *Mathematics Subject Classification* (1985 Revision). Primary 47A70; Secondary 47A67.
Key words and phrases. Moment problem, closed operator.

b) Suppose that U is minimal and denote by (u^n) its biorthogonal system. Then the images of

$$\sum_{n=1}^{k} c_n u^n \qquad (k = 1, 2, \ldots)$$

are dense in l_2. Conversely, suppose that $\overline{\operatorname{Im} B} = l_2$ and consider a sequence $c=(c_n)\in l_2$ such that $\sum c_n u_n = 0$. For any $x\in D$ we have

$$\langle Bx, c\rangle = \sum \langle x, u_n\rangle \overline{c_n} = \langle x, \sum c_n u_n\rangle = 0$$

which implies $c=0$.

c) Let $U=U_0\cup U_1$, where $U_0=(u_n^0)$ is a Riesz basis in H and let $U_1=(u_n^1)_{n=1}^{\omega}$ $\omega\leqq\infty$. Denote by (b_n^0) the system biorthogonal to U_0 and consider the expansions

$$u_j^1 = \sum_{n=1}^{\infty} \langle u_j^1, b_n^0\rangle u_n^0.$$

$$\beta_j := (\langle u_j^1, b_n^0\rangle)_{e=1}^{\infty} \in l_2 \quad \text{and define} \quad \Omega := l_2$$

for $\omega=\infty$ and $\Omega:=\mathbf{C}^n$ for $\omega<\infty$. Then for any $f\in H$, $f=\sum c_n b_n^0$ we have $Bf=$ $=\{c, Ac\}$, where $A\colon l_2\to\Omega$ is defined by the formula

$$Ac := (\langle c, \beta_j\rangle)_{j=1}^{\omega}.$$

As every moment operator, A is closed, hence $\operatorname{Im} B = \operatorname{Graf} A$ is closed. Consider the standard basis in Ω: $\delta_j := (0, \ldots, \overset{j}{\breve{1}}, 0, \ldots)$ if $\omega=\infty$ and $\delta_j := (0, \ldots, \overset{j}{\breve{1}}, 0, \ldots, \overset{\omega}{\breve{0}})$ if $\omega<\infty$. Then $\{\beta_j, -\delta_j\}\perp \operatorname{Im} B$ and they are obviously linearly independent. Thus we have $\omega\leqq \operatorname{codim} \operatorname{Im} B$, and it remains only to show that in case $\omega<\infty$ the vectors $\{\beta_j, -\delta_j\}_{j=1}^{\omega}$ span the orthocomplement of $\operatorname{Im} B$. Indeed,

$$\{c', d\}\perp \operatorname{Im} B \Leftrightarrow 0 = \langle\{c, Ac\}, \{c', d\}\rangle = \langle c, c'\rangle + \langle Ac, d\rangle =$$

$$= \langle c, c' + A^* d\rangle \quad \text{for all} \quad c\in l_2 \Leftrightarrow c' + A^* d = c' + \sum_{j=1}^{\omega} \beta_j d_j = 0$$

hence

$$(\operatorname{Im} B)^{\perp} = \Big\{\Big\{\sum_{j=1}^{\omega} \beta_j d_j, -d\Big\} : d\in \mathbf{C} = \Omega\Big\}.$$

The proof of the theorem is complete.

REFERENCES

[1] А. Г. Бутковский, Применение некоторых результатов теории чисел в проблеме финитного управления и управляемости в распределенных системах, *Dokl. Akad. Nauk SSSR* **227** (1976), 309—311.

[2] Bari, N. K., Sur les systèmes complets de fonctions orthogonales, *Rec. Math.* [*Mat. Sbornik*] N. S. **14** (56) (1944), 51—108. *MR* **6**—267.

[3] Joó, I., On the vibration of a string, *Studia Sci. Math. Hungar.* **22** (1987), 1—9.

[4] Joó, I., On the reachability set of a string, *Acta Math. Hungar.* **49** (1987), 203—211.

[5] Horváth, M., Vibrating strings with free ends, *Acta Math. Hungar.* **51** (1988), 171—180.

[6] Avdonin, S. A. and Ivanov, S. A., Serial bases of exponentials and the problem of damping of a system of strings, *Dokl. Akad. Nauk SSSR* **275** (1984), No 2, 355—358 (in Russian). *MR* **86d**: 35078.

[7] Nikol'skiĭ, N. K., *Treatise on the shift operator*, Springer-Verlag, Berlin, 1986.

(Received September 1, 1986)

EÖTVÖS LORÁND TUDOMÁNYEGYETEM
TERMÉSZETTUDOMÁNYI KAR
ALKALMAZOTT ANALÍZIS TANSZÉK
POSTAFIÓK 323
H—1445 BUDAPEST 8
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 263—268

ON A LATTICE-POINT PROBLEM OF H. STEINHAUS

JÓZSEF BECK

Many years ago Hugo Steinhaus proposed the following problem (see for example Sierpiński [3] and Moser [2] Problem 59).

Does there exist a point set such that no matter how it is placed on the plane it covers exactly one integer lattice point (*Steinhaus' property*, in short)?

Sierpiński [3] remarked that there is no bounded and closed or bounded and open set satisfying Steinhaus' property above. The aim of this note is to prove the following much stronger statement.

THEOREM. *There is no bounded and Lebesgue measurable set satisfying Steinhaus' property.*

PROOF. Let μ denote the two-dimensional Lebesgue measure (area). Let $S\subset \mathbf{R}^2$ be a Lebesgue measurable set with $\mu(S)<\infty$. Observe that (card stand for cardinality)

$$\int\limits_{[0,1)^2} \operatorname{card}\left((S+\mathbf{x}))\cap \mathbf{Z}^2\right) d\mathbf{x} = \mu(S). \tag{1}$$

Suppose now that $S\subset \mathbf{R}^2$ satisfies Steinhaus' property. If $\mu(S)>1$ (or $\mu(S)<1$) then by (1) there exists a vector $\mathbf{x}_1\in[0, 1)^2$ (or $\mathbf{x}_2\in[0, 1)^2$) such that

$$\operatorname{card}(S+\mathbf{x}_1)\cap\mathbf{Z}^2 \geqq 2 \quad (\text{or } \operatorname{card}(S+\mathbf{x}_2)\cap\mathbf{Z}^2 = 0),$$

which contradicts Steinhaus' property.

We may therefore assume that $\mu(S)=1$. Assume further that S is bounded.

If S is closed, then we are done as follows: One can find Lebesgue measurable sets $A_1\supset A_2\supset A_3\supset\ldots$ such that $S=\bigcap\limits_{i=1}^{\infty} A_i$ and $\mu(A_i)=1+\frac{1}{i}$.

It follows from (1) by a standard compactness argument that there exist an infinite subsequence $A_{i_1}\subset A_{i_2}\subset A_{i_3}\supset\ldots$, vectors $\mathbf{x}_{i_1}, \mathbf{x}_{i_2}, \mathbf{x}_{i_3}, \ldots$ in $[0, 1)^2$ and two distinct lattice points $\{\mathbf{n}, \mathbf{m}\}\subset \mathbf{Z}^2$ such that

$$\{\mathbf{n}, \mathbf{m}\} \subset A_{i_j}+\mathbf{x}_{i_j}, \quad j\geqq 1 \quad \text{and} \quad \lim_{j\to\infty} \mathbf{x}_{i_j} = \mathbf{x}_0 \quad \text{exists.}$$

1980 *Mathematics Subject Classification* (1985 Revision). Primary 10E99; Secondary 10J25.
Key words and phrases. Lattice-points, Fourier analysis, Jensen's formula.

Since S is closed, we conclude that

$$\{\mathbf{n}, \mathbf{m}\} \subset S+\mathbf{x}_0,$$

which contradicts Steinhaus' property.

The same argument shows that if S is bounded and $\mu(S)\geqq 1$, then there is a vector $\mathbf{x}\in[0, 1)^2$ and a pair $\{\mathbf{n}, \mathbf{m}\}$ of lattice points such that $\{\mathbf{n}, \mathbf{m}\}\subset \mathrm{cl}\,(S)+\mathbf{x}$ (cl stands for the *closure* operation). Since S satisfies Steinhaus' property, we may assume that (say) $\mathbf{n}\notin S$. If S is open, after a sufficiently small translation $\mathbf{y}$ we have

$$\{\mathbf{n}, \mathbf{m}\} \subset S+\mathbf{x}+\mathbf{y},$$

again a contradiction.

In contrast to these simple particular cases, in the general case we need *rotation*. Indeed, let for example $S=[0, 1)^2$; obviously any translate $[0, 1)^2+\mathbf{x}$ of the half-open unit square contains exactly one lattice point.

We now begin the proof of the general case. The proof is based on a Fourier transform approach. Let $S\subset \mathbf{R}^2$ be a bounded Lebesgue measurable set satisfying Steinhaus' property and $\mu(S)=1$. Without loss of generality, we may assume that $0\in S$. Let $D=D(S)$ denote the diameter of S. Let $N\geqq 10D(S)$ be an integral parameter to be specified later.

For any Lebesgue measurable set $A\subset \mathbf{R}^2$, let

$$\mu_0(A) = \mu\left(A\cap\left[-N-\frac{1}{2},\ N+\frac{1}{2}\right]^2\right).$$

For any bounded set $A\subset \mathbf{R}^2$, let

$$Z_0(A) = \mathrm{card}\, A\cap\left[-N-\frac{1}{2},\ N+\frac{1}{2}\right]^2\cap \mathbf{Z}^2.$$

For any *rotation* (angle) $\tau\in[0, 2\pi)$ and *translation* $\mathbf{x}\in \mathbf{R}^2$, write

$$S(\tau, \mathbf{x}) = \{\tau\mathbf{y}+\mathbf{x}\colon\ \mathbf{y}\in S\}.$$

Let χ_τ denote the characteristic function of the set $S(\tau, 0)$. Let

$$f_\tau = \chi_\tau*(dZ_0-d\mu_0) \tag{2}$$

where $*$ stands for the *convolution* operation. More explicitly, we have

$$f_\tau(\mathbf{x}) = \int_{\mathbf{R}^2} \chi_\tau(\mathbf{x}-\mathbf{y})\, dZ_0(\mathbf{y}) - \int_{\mathbf{R}^2} \chi_\tau(\mathbf{x}-\mathbf{y})\, d\mu_0(\mathbf{y}) =$$

$$= \mathrm{card}\left(S(\tau, -\mathbf{x})\cap\left[-N-\frac{1}{2},\ N+\frac{1}{2}\right]^2\cap \mathbf{Z}^2\right)-\mu\left(S(\tau, -\mathbf{x})\cap\left[-N-\frac{1}{2},\ N+\frac{1}{2}\right]^2\right).$$

Since S satisfies Steinhaus' property and $\mu(S)=1$, we have

$$f_\tau(\mathbf{x}) = 0 \quad \text{provided} \quad S(\tau, -\mathbf{x}) \subset \left[-N-\frac{1}{2},\ N+\frac{1}{2}\right]^2, \tag{3}$$

and

(4) $$|f_\tau(\mathbf{x})| \leqq 1 \quad \text{provided} \quad S(\tau, -\mathbf{x}) \cap \left[-N-\frac{1}{2}, N+\frac{1}{2}\right]^2 \neq \emptyset.$$

Moreover, evidently

(5) $$f_\tau(\mathbf{x}) = 0 \quad \text{provided} \quad S(\tau, -\mathbf{x}) \cap \left[-N-\frac{1}{2}, N+\frac{1}{2}\right]^2 = \emptyset.$$

Combining (3)—(5), we obtain

(6) $$\int_{\mathbf{R}^2} (f_\tau(\mathbf{x}))^2 \, d\mathbf{x} \leqq 8D(2N+1) \quad \text{for all} \quad \tau \in [0, 2\pi).$$

We recall some well-known facts from Fourier analysis. Given any function $F \in L^2(\mathbf{R}^K)$, let

$$\hat{F}(\mathbf{t}) = (2\pi)^{-K/2} \int_{\mathbf{R}^K} e^{-i\mathbf{t}\cdot\mathbf{x}} F(\mathbf{x}) \, d\mathbf{x}$$

denote the *Fourier transform* of F. Here $i = \sqrt{-1}$ and $\mathbf{x} \cdot \mathbf{t} = x_1 t_1 + x_2 t_2 + \ldots + x_K t_K$. It is known

(7) $$\widehat{F * G} = \hat{F} \cdot \hat{G};$$

(8) $$\int_{\mathbf{R}^K} |F(\mathbf{x})|^2 \, d\mathbf{x} = \int_{\mathbf{R}^K} |\hat{F}(\mathbf{t})|^2 \, d\mathbf{t} \qquad \text{(Plancherel identity)}$$

where $F, G \in L^2(\mathbf{R}^K)$ and $*$ stands for the convolution operation.

By (2), (7) and (8) we have

(9) $$\int_{\mathbf{R}^2} (f_\tau(\mathbf{x}))^2 \, d\mathbf{x} = \int_{\mathbf{R}^2} |\hat{\chi}_\tau(\mathbf{t})|^2 |\varphi(\mathbf{t})|^2 \, d\mathbf{t}$$

where

(10) $$\varphi(\mathbf{t}) = (d\hat{Z}_0 - d\hat{\mu}_0)(\mathbf{t}) = \frac{1}{2\pi} \int_{\mathbf{R}^2} e^{-i\mathbf{t}\cdot\mathbf{x}} (dZ_0(\mathbf{x}) - d\mu_0(\mathbf{x})).$$

An elementary calculation shows that $(\mathbf{t} = (t_1, t_2))$

(11) $$\varphi(\mathbf{t}) = \frac{1}{2\pi} \sin\left(\frac{2N+1}{2} t_1\right) \sin\left(\frac{2N+1}{2} t_2\right) \left(\frac{1}{\sin\frac{t_1}{2} \cdot \sin\frac{t_2}{2}} - \frac{1}{\frac{t_1}{2} \cdot \frac{t_2}{2}}\right).$$

We need the following lemma.

LEMMA. *There exist* $\theta \in [0, 2\pi)$ *and* $\mathbf{n} \in \mathbf{Z}^2 \setminus \{\mathbf{0}\}$ *such that* $\hat{\chi}_\theta(2\pi\mathbf{n}) \neq 0$.

First we derive the theorem from the lemma. We specify the value of the parameter N: suppose that N is sufficiently large such that the inequality

(12) $$2N+1 > \max\left\{\frac{2\sqrt{2}D}{|\hat{\chi}_\theta(2\pi\mathbf{n})|}, \frac{32 \cdot 10^4 \cdot D}{(2\pi)^2 |\hat{\chi}_\theta(2\pi\mathbf{n})|^2}\right\}$$

holds. Using the elementary inequality $|x| \geqq |\sin x| \geqq \frac{2}{\pi}|x|$ for $|x| \leqq \frac{\pi}{2}$, by (11) we have with $\mathbf{n}=(n_1, n_2)$ and $\mathbf{t}=(t_1, t_2)$,

$$\varphi(\mathbf{t}) \geqq \frac{1}{2\pi}\left(\frac{2}{\pi}(2N+1)\right)^2 - \frac{1}{2\pi}(2N+1)\frac{1}{\pi} > \frac{(2N+1)^2}{100} \tag{13}$$

whenever

$$|t_1 - 2\pi n_1| \leqq \frac{\pi}{2N+1} \quad \text{and} \quad |t_2 - 2\pi n_2| \leqq \frac{\pi}{2N+1}.$$

Furthermore, using $\mathbf{0} \in S$ and $|e^{iy} - e^{iz}| \leqq |y-z|$, y, z real numbers, we have

$$|\hat{\chi}_\theta(\mathbf{t}) - \hat{\chi}_\theta(\mathbf{u})| = \left|\frac{1}{2\pi} \int_{S(\theta,0)} (e^{-i\mathbf{t}\cdot\mathbf{x}} - e^{-i\mathbf{u}\cdot\mathbf{x}})\, d\mathbf{x}\right| \leqq \frac{D|\mathbf{t}-\mathbf{u}|}{2\pi}\mu(S) = \frac{D|\mathbf{t}-\mathbf{u}|}{2\pi} \tag{14}$$

where $|\mathbf{x}| = (x_1^2 + x_2^2)^{1/2}$ stands for the usual euclidean distance. Let $Q(y) = [-y, y]^2$ ($y > 0$ real number). It follows from (12)—(14) that

$$\int_{Q(\pi/2N+1)+2\pi\mathbf{n}} |\hat{\chi}_\theta(\mathbf{t})|^2 |\varphi(\mathbf{t})|^2\, d\mathbf{t} \geqq \Big(\min_{\mathbf{t} \in Q(\pi/2N+1)+2\pi\mathbf{n}} |\hat{\chi}_\vartheta(\mathbf{t})|^2\Big) \int_{Q(\pi/2N+1)+2\pi\mathbf{n}} |\varphi(\mathbf{t})|^2\, d\mathbf{t} \geqq$$

$$\geqq \left(\frac{1}{2}|\hat{\chi}_\theta(2\pi\mathbf{n})|\right)^2 \left(\frac{(2N+1)^2}{100}\right)^2 \left(\frac{2}{2N+1}\right)^2 > 8D(2N+1).$$

Combining this with (9) we conclude that

$$\int_{\mathbf{R}^2} (f_\theta(\mathbf{x}))^2\, d\mathbf{x} > 8D(2N+1),$$

which contradicts (6). This proves the theorem. It remains to prove the lemma.

PROOF of the lemma. The proof will be by contradiction. Suppose that $\hat{\chi}_\tau(2\pi\mathbf{m}) = 0$ for all $\tau \in [0, 2\pi)$ and $\mathbf{m} \in \mathbf{Z}^2 \setminus \{\mathbf{0}\}$. Let χ_S denote the characteristic function of the set S. We have

$$\hat{\chi}_S(\mathbf{t}) = \frac{1}{2\pi}\int_S e^{-i\mathbf{t}\cdot\mathbf{x}}\, d\mathbf{x} = \frac{1}{(2\pi)^{1/2}} \frac{1}{(2\pi)^{1/2}} \int_{\mathbf{R}} e^{-i|\mathbf{t}|x} g_\beta(x)\, dx \tag{15}$$

where $\beta = \beta(\mathbf{t}) \in [0, 2\pi)$ is defined by the equation

$$\mathbf{t} = (|\mathbf{t}| \cos\beta, |\mathbf{t}| \sin\beta)$$

and $g_\beta(x)$ is the one-dimensional Lebesgue measure of the set

$$\{\mathbf{x} \in S\colon\ \mathbf{x}\cdot\mathbf{t} = x \cdot |\mathbf{t}|\}$$

(by Fubini's theorem, $g_\beta(x)$ is well defined for almost all $x \in \mathbf{R}$). By (15),

$$\hat{\chi}_S(\mathbf{t}) = \frac{1}{(2\pi)^{1/2}}\, \hat{g}_\beta(|\mathbf{t}|) \tag{16}$$

where $\hat{g}_\beta$ denotes the Fourier transform of the function $g_\beta(x)$ of one variable $x \in \mathbf{R}$.

Using the trivial identity $\hat{\chi}_\tau(\mathbf{u})=\hat{\chi}_S(\tau^{-1}\mathbf{u})$, $\mathbf{u}\in\mathbf{R}^2$, by (16) we have with $\beta=\beta(\mathbf{t})$,

$$\hat{\chi}_\beta(\mathbf{t}) = \hat{\chi}_S(\beta^{-1}\,\mathbf{t}) = \frac{1}{(2\pi)^{1/2}}\,\hat{g}_0(|\mathbf{t}|) \tag{17}$$

where

$$\hat{g}_0(t) = \frac{1}{(2\pi)^{1/2}}\int_{\mathbf{R}} e^{-itx}\,g_0(x)\,dx$$

and $g_0(x)$ denotes the one-dimensional Lebesgue measure of the set $\{\mathbf{x}=(x_1,x_2)\in S: x_1=x\}$. By the hypothesis of the lemma and (17), we obtain for all $\mathbf{m}=(m_1,m_2)\in \in\mathbf{Z}^2\setminus\{\mathbf{0}\}$,

$$0 = (2\pi)^{1/2}\,\hat{g}_0(2\pi\,|\mathbf{m}|) = \int_{\mathbf{R}} e^{-i2\pi|\mathbf{m}|x} g_0(x)\,dx \tag{18}$$

where

$$|\mathbf{m}| = (m_1^2+m_2^2)^{1/2}.$$

For any complex number $z\in\mathbf{C}$, write

$$G(z) = \int_{\mathbf{R}} e^{-izx}\,g_0(x)\,dx.$$

By (18) we have for all $\mathbf{m}\in\mathbf{Z}^2\setminus\{\mathbf{0}\}$,

$$G(2\pi\,|\mathbf{m}|) = 0. \tag{19}$$

Since $\mathbf{0}\in S$, $g_0(x)=0$ for all $|x|>D=D(S)$. Hence

$$|G(z)| \leqq e^{D|z|}\int_{\mathbf{R}} g_0(x)\,dx = e^{D|z|}\mu(S) = e^{D|z|}. \tag{20}$$

Applying Jensen's formula (see for example Ahlfors [1]) to the entire function $G(z)$, $z\in\mathbf{C}$, by (19) we obtain for any real number $r>0$,

$$\log\big(\max_{|z|=r}|G(z)|\big) \leqq \frac{1}{2\pi}\int_0^{2\pi}\log|G(re^{i\theta})|\,d\theta \geqq \log|G(0)| + \sum_{\substack{\mathbf{m}\in\mathbf{Z}^2\setminus\{0\}:\\ |\mathbf{m}|\leqq(r/2\pi)}} \log\left(\frac{r}{2\pi\,|\mathbf{m}|}\right). \tag{21}$$

We have

$$G(0) = \int_{\mathbf{R}} g_0(x)\,dx = \mu(S) = 1, \tag{22}$$

and

$$\sum_{\substack{\mathbf{m}\in\mathbf{Z}^2\setminus\{0\}:\\ |\mathbf{m}|\leqq(r/2\pi)}} \log\left(\frac{r}{2\pi\,|\mathbf{m}|}\right) \geqq \sum_{\substack{\mathbf{m}\in\mathbf{Z}^2\setminus\{0\}:\\ |\mathbf{m}|\leqq(r/2\pi e)}} \log\left(\frac{r}{2\pi\,|\mathbf{m}|}\right) \geqq \sum_{\substack{\mathbf{m}\in\mathbf{Z}^2\setminus\{0\}:\\ |\mathbf{m}|\leqq(r/2\pi e)}} 1 \geqq \pi\left(\frac{r}{2\pi e}\right)^2 - O(r). \tag{23}$$

Combining (20)—(23), we conclude that

$$Dr \geqq \pi \left(\frac{r}{2\pi e}\right)^2 - O(r).$$

But this is a contradiction if $r \to \infty$, and the lemma follows.

ACKNOWLEDGEMENT. The author is very indebted to Prof. G. Halász who has simplified the proof of the lemma.

REFERENCES

[1] AHLFORS, L. V., *Complex Analysis. An introduction to the theory of analytic functions of one complex variable,* McGraw-Hill, New York—Toronto—London, 1953, 1966 and 1979. *MR* **14**—857.

[2] MOSER, W., *Research problems in discrete geometry,* Mimeograph notes, 1981.

[3] SIERPIŃSKI, W., Sur un problème de H. Steinhaus concernant les ensembles de points sur le plan, *Fund. Math.* **46** (1959), 191—194. *MR* **21** #851.

(Received September 2, 1986)

EÖTVÖS LORÁND TUDOMÁNYEGYETEM
TERMÉSZETTUDOMÁNYI KAR
SZÁMITÓGÉPTUDOMÁNYI TANSZÉK
MÚZEUM KRT. 6—8.
H—1088 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 269—275

ON THE SPEED OF CONVERGENCE FOR CRITICAL GALTON—WATSON PROCESSES

N. ARATÓ

1. Introduction

Let $\mu(k)$, $k=0, 1, 2, \ldots$, denote the number of individuals in the k-th generation of a critical Galton—Watson process with basic random variable X ($EX=1$, $D^2X=\sigma^2$). (The definitions see in [10].) That is let $\mu(0)=b>0$ and set $\mu(k)=\sum_{i=1}^{\mu(k-1)} X_{k,i}$, where $X_{k,i}$, $k, i=1, 2, \ldots$, are independent copies of X. From the first n generation let us define the continuous time processes in the following way

$$Y_n(t)=\frac{\mu([nt])-b}{\sigma\sqrt{nb}}, \quad t\in[0, 1].$$

In this work using the embedding from [6, 7] I will examine how big $|Y_n(t)-W(t)|$ can be, where $w(t)$ means a standard Wiener process.

The first results on the convergence of a sequence of normalized Galton—Watson processes can be found in the works of Feller [3], Lamperti [8] and Jirina [5].

It was shown by Lindwall [9] that if $\mu^{(n)}(0)=b_n$ and $\frac{b_n}{n}\to\infty$, then the process $\frac{\mu^{(n)}([nt])-b_n}{\sigma\sqrt{nb_n}}$ converges weakly in $D[0, 1]$ to a standard Wiener process, and if $b_n=n$, then $\frac{\mu^{(n)}([nt])}{n}$ converges weakly to a Feller process in $D[0, 1]$.

In [4] Grimwall studied the convergence, when the processes are not necessarily critical, but nearly critical.

K. A. Borovkov in [1] estimated the speed of convergence to the Feller process in the "nearly critical" case. It is not easily seen what is the relation between his results and the present ones.

1980 *Mathematics Subject Classification* (1985 Revision). Primary 60F17; Secondary 60J80, 60J65.

Key words and phrases. Weak convergence, branching process, Wiener process, Galton—Watson process.

2. The main theorems

Let $\mu(k)$, X, b, $X_{k,i}$ and Y_n have the same meaning as in the introduction. I will prove the next two theorems.

THEOREM 1. *Assume that* $\mathsf{E}|X-1|^{2+\gamma}=D(\gamma)<\infty$ *for some* $\gamma>0$. *Then there exists a probability space with a* $\mu(k)$ *and a* $w(t)$ *Wiener process on it, that for every* n

$$(1)\qquad \mathsf{P}(\sup_{0\le t\le 1}|Y_n(t)-w(t)|>\varepsilon)<u_1 ne^{-(\varepsilon^2 n/12)}+\frac{c(\gamma)D(\gamma)}{\varepsilon^{4+2\gamma}}\left(\frac{\log\frac{b}{n}}{\sqrt{\frac{b}{n}}}\right)^{2+\gamma}+$$
$$+u_2 e^{-u_3\varepsilon^2\sqrt{b/n}\log(b/n)}+\frac{n}{b}+A_1\frac{1}{A_2^{2+\gamma}}\frac{1}{\varepsilon^{2+\gamma}}\frac{1}{(nb)^{\gamma/2}}+u_4\frac{1}{\varepsilon^2}\left(\log\frac{b}{n}\right)\left(\frac{n}{b}\right)^4$$

if

$$8(nb)^{(1/2+\gamma)-(1/2)}<\varepsilon<\min\left\{14\left(\log\frac{b}{n}\right)^{1/2},\ A_3(nb\log nb)^{1/2}\right\},$$

where u_i $(i=1,\dots,4)$ *are absolute constants,* $c(\gamma)$ *depends only on* γ, *and* A_i $(i=1,2,3)$ *depend only on the distribution of* X.

THEOREM 2. *Assume that* $\mathsf{E}e^{tX}$ *exists in a neighbourhood of* $t=0$. *Then there exists a probability space with* $\mu(k)$ *and a* $w(t)$ *Wiener process on it such that*

$$\mathsf{P}(\sup_{0\le t\le 1}|Y_n(t)-w(t)|>\varepsilon)\le V_1 ne^{-(\varepsilon^2 n/12)}+V_2\left(\log\frac{b}{n}\right)\left(\frac{n}{b}\right)^3\frac{1}{\varepsilon^2}+V_3e^{-V_4\varepsilon^2\sqrt{b/n}/\log(b/n)}+$$
$$+V_5e^{-V_4\varepsilon^2(b/n)/\log^2(b/n)}+\frac{n}{b}+Ke^{-\mu(V_7\varepsilon^2 b/\log(b/n)-C\log nb)}+Ke^{-\mu(1/4)\varepsilon\sqrt{bn}-C\log nb)}$$

holds true for every n, *if* $\varepsilon\le 12\left(\log\frac{b}{n}\right)^{1/2}$, *where* V_i $(i=1,\dots,7)$ *are positive absolute constants and* K,μ,C *depend only on the distribution function of* X.

PROOF of Theorem 1. To simplify the proof let us assume that $DX=\sigma=1$.

Let Z_i, $i=1,2,3,\dots$, be independent copies of $(X-1)$. One can take the random variables $X_{k,i}$ successively from the sequence Z_i+1, $i=1,2,\dots$. Then

$$(2)\quad \mu(k)-b=\sum_{i=1}^{\mu(k-1)}X_{k,i}-b=\sum_{i=1}^{\mu(k-1)}(X_{k,i}-1)+\mu(k-1)-b=\sum_{i=1}^{\mu(0)+\dots+\mu(k-1)}Z_i.$$

On the basis of Theorem 4 in [7] one can conclude that there exists a probability space and a Wiener process $\tilde{w}(t)$ on it such that

$$(3)\qquad \mathsf{P}(\max_{1\le k\le n}|S_k-\tilde{w}(k)|>x)\le B_1\frac{1}{B_2^{2+\gamma}}\frac{n}{x^{2+\gamma}}$$

holds true for every n and any x such that

$$(4)\qquad n^{1/2+\gamma}<x<B_3(n\log n)^{1/2},$$

where $S_k=\sum_{i=1}^{k} Z_i$ and B_l $(l=1, 2, 3)$ are positive constants depending only on the distribution function of X.

Set $W(t)=\frac{1}{\sqrt{nb}}\tilde{W}(tnb)$. For all positive a we have

$$\mathsf{P}\left(\max_{0\leq k\leq n}\left|\frac{\mu(k)-b}{\sqrt{nb}}-w\left(\frac{k}{n}\right)\right|>\lambda\right)=\mathsf{P}\left(\max_{0\leq k\leq n}\left|\mu(k)-b-\sqrt{nb}\,w\left(\frac{k}{n}\right)\right|>\lambda\sqrt{nb}\right)=$$

$$\mathsf{P}\big(\max_{0\leq k\leq n}|\mu(k)-b-\tilde{w}(\mu(0)+\ldots+\mu(k-1))+\tilde{w}(\mu(0)+\ldots+\mu(k-1))-\tilde{w}(kb)>$$

$$>\lambda\sqrt{nb}\big)\leq\mathsf{P}\big(\max_{0\leq k\leq n}|\mu(0)+\ldots+\mu(k-1)-kb|>a\big)+\mathsf{P}\big(\max_{0\leq k\leq n}|\mu(0)+\ldots$$

$$\ldots+\mu(k-1)-kb|\leq a,\ \max_{0\leq k\leq n}|\mu(k)-b-\tilde{w}(\mu(0)+\ldots+\mu(k-1))+\tilde{w}(\mu(0)+\ldots$$

$$\ldots+\mu(k-1))-\tilde{w}(kb)\big|>\lambda\sqrt{nb}\big).$$

From this inequality and from (2) it is easy to see that

(5)
$$\mathsf{P}\left(\max_{0\leq k\leq n}\left|\frac{\mu(k)-b}{\sqrt{nb}}-w\left(\frac{k}{n}\right)\right|>\lambda\right)\leq\mathsf{P}\big(\max_{0\leq k\leq n}|\mu(0)+\ldots+\mu(k-1)-kb|>a\big)+$$

$$+\mathsf{P}\left(\max_{0\leq l\leq nb+a}|S_l-\tilde{w}(l)|>\frac{1}{2}\lambda\sqrt{nb}\right)+\mathsf{P}\left(\max_{0\leq l\leq nb}\max_{0\leq m\leq a}|\tilde{w}(l+m)-\tilde{w}(l)|>\frac{1}{2}\lambda\sqrt{nb}\right).$$

Now first I estimate $\mathsf{P}\big(\max_{0\leq k\leq n}|\mu(0)+\ldots+\mu(k-1)-kb|>a\big)$ with the help of a martingal inequality getting an upper bound for $\mathsf{P}\left(\max_{0\leq k\leq n}\left|\frac{\mu(k)-b}{\sqrt{nb}}-w\left(\frac{b}{n}\right)\right|>\lambda\right)$. Applying this upper bound we can improve the estimator for

$$\mathsf{P}\big(\max_{0\leq k\leq n}|\mu(0)+\ldots+\mu(k-1)-kb|>a\big).$$

Being a critical Galton—Watson process $\mu(k)$ is a martingale and the Kolmogorov—Doob inequality applies

$$\mathsf{P}\big(\max_{0\leq k\leq n}|\mu(0)+\ldots+\mu(k-1)-kb|>a\big)\leq$$

(6)
$$\leq\mathsf{P}\left(\max_{0\leq k\leq n-1}|\mu(k)-b|>\frac{a}{n}\right)\leq\frac{\mathsf{D}^2\mu(n-1)}{(a/n)^2}\leq\frac{n^3b}{a^2}.$$

It is obvious that

$$\mathsf{P}\left(\max_{0\leq l\leq nb+a}|S_l-\tilde{w}(l)|>\frac{1}{2}\lambda\sqrt{nb}\right)\leq$$

(7)
$$\leq\mathsf{P}\left(\max_{0\leq l\leq nb+a}|S_l|>\frac{1}{4}\lambda\sqrt{nb}\right)+\mathsf{P}\left(\max_{0\leq l\leq nb+a}|\tilde{w}(l)|>\frac{1}{4}\lambda\sqrt{nb}\right),$$

and

$$\text{(8)}\quad \mathsf{P}\left(\max_{0\le l\le nb}\ \max_{0\le m\le a}|\tilde{w}(l+m)-\tilde{w}(l)|>\frac{1}{2}\lambda\sqrt{nb}\right)\le \mathsf{P}\left(\max_{0\le l\le nb+a}|\tilde{w}(l)|>\frac{1}{4}\lambda\sqrt{nb}\right).$$

It is known that $E(|S_n|^{2+\gamma})\le C_1(2+\gamma)\mathsf{E}(|Z_1|^{2+\gamma})n^{2+\gamma/2}$, where $C_1(2+\gamma)$ is a positive constant depending only on γ. Thus the Kolmogorov inequality together with (7) implies that

$$\mathsf{P}\left(\max_{0\le l\le nb+a}|S_l|>\frac{1}{4}\lambda\sqrt{nb}\right)<\frac{C_1(2+\gamma)\,\mathsf{D}(\gamma)(nb+a)^{2+\gamma/2}}{\left(\frac{1}{4}\lambda\sqrt{nb}\right)^{2+\gamma}}+\varkappa_1 e^{(-1/4)\,\lambda\sqrt{nb/nb+a}},$$

where $\varkappa_2$ is an absolute constant. Putting the pieces together with $a=nb$

$$\text{(9)}\quad \mathsf{P}\left(\max_{0\le k\le n}\left|\frac{\mu(k)-b}{\sqrt{nb}}-w\left(\frac{k}{n}\right)\right|>\lambda\right)\le\frac{C_2(2+\gamma)\,\mathsf{D}(2+\gamma)}{\lambda^{2+\gamma}}+\varkappa_2 e^{-\varkappa_3\lambda}+\frac{n}{b},$$

where $\varkappa_2$, $\varkappa_3$ are positive absolute constants and $C_2(2+\gamma)$ depends only on γ. Further, it is clear that

$$\text{(10)}\quad \begin{aligned}&\mathsf{P}\left(\max_{0\le k\le n}|\mu(0)+\ldots+\mu(k-1)-kb|>a\right)\le\\ &\le \mathsf{P}\left(\max_{0\le k\le n}\left|\frac{\mu(k)-b}{nb}-w\left(\frac{k}{n}\right)\right|>\frac{1}{2}\frac{a}{n\sqrt{nb}}\right)+\mathsf{P}\left(\sup_{0\le t\le 1}|w(t)|>\frac{1}{2}\frac{a}{n\sqrt{nb}}\right).\end{aligned}$$

Hence by (9) we have

$$\text{(11)}\quad \mathsf{P}\left(\max_{0\le k\le n}|\mu(0)+\ldots+\mu(k-1)-kb|>a\right)\le\frac{C_3(2+\gamma)\,\mathsf{D}(\gamma)}{(a/n\sqrt{nb})^{2+\gamma}}+\varkappa_4 e^{-\varkappa_5(a/n\sqrt{nb})}+\frac{n}{b}.$$

The following inequality can be used (see [2] page 29):

$$\text{(12)}\quad \mathsf{P}\left(\sup_{0\le s\le T-h}\ \sup_{0\le t\le h}|w(s+t)-w(s)|>v\sqrt{h}\right)<\frac{\varkappa_6 T}{h}e^{-(v/3)^2},$$

where $\varkappa_6$ is an absolute constant. Now we have

$$\text{(13)}\quad \mathsf{P}\left(\max_{0\le l\le nb}\ \max_{0\le m\le a}|\tilde{w}(l+m)-\tilde{w}(l)|>\frac{1}{2}\lambda\sqrt{nb}\right)<\varkappa_6\frac{nb+ae^{-(\lambda^2 nb/12a)}}{a}.$$

Now let us return to the inequality (5) applying inequalities (3), (11) and (13) we get:

$$\begin{aligned}\mathsf{P}\left(\max_{0\le k\le n}\left|\frac{\mu(k)-b}{\sqrt{nb}}-w\left(\frac{k}{n}\right)\right|>\lambda\right)&\le\frac{C_3(2+\gamma)\,\mathsf{D}(\gamma)}{(a/n\sqrt{nb})^{2+\gamma}}+\varkappa_4 e^{-\varkappa_5(a/n\sqrt{nb}}+\frac{n}{b}+\\ &+B_1\frac{1}{B_2^{\frac{2}{2}+\gamma}}\frac{nb+a}{\left(\frac{1}{2}\lambda\sqrt{nb}\right)^{2+\gamma}}+\varkappa_6\frac{nb+a}{a}e^{-(\lambda^2 nb/12a)},\end{aligned}$$

if

$$(nb+a)^{(1/2+\gamma)} < \frac{1}{2}\lambda\sqrt{nb} < B_3((nb+a)\log(nb+a))^{1/2}.$$

This inequality for $a=\dfrac{\lambda^2 nb}{48\log\dfrac{b}{n}}$ implies

$$\text{(14)}\qquad \mathsf{P}\left(\max_{0\le k\le n}\left|\frac{\mu(k)-b}{\sqrt{nb}}-w\left(\frac{k}{n}\right)\right|>\lambda\right) \le \frac{c_4(2+\gamma)\,\mathsf{D}(\gamma)}{\lambda^{4+2\gamma}}\left(\frac{\log\dfrac{b}{n}}{\sqrt{\dfrac{b}{n}}}\right)^{2+\gamma}+$$

$$+\varkappa_4 e^{-\varkappa_7\lambda^2\sqrt{b/n}/\log(b/n)}+\frac{n}{b}+B_4\frac{1}{B_5^{2+\gamma}}\frac{1}{\lambda^{2+\gamma}}\frac{1}{(nb)^{\gamma/2}}+\varkappa_8\frac{1}{\lambda^2}\left(\log\frac{b}{n}\right)\left(\frac{n}{b}\right)^4,$$

if

$$\text{(15)}\qquad 4(nb)^{(1/2+\gamma)-(1/2)} < \lambda < \min\left\{7\left(\log\frac{b}{n}\right)^{1/2},\ B_3(nb\log nb)^{1/2}\right\},$$

where $\varkappa_7$, $\varkappa_8$ are absolute constants, $C_4(2+\gamma)$ depends only on γ, B_4 and B_5 depend only on the distribution of X.

On the other side it is obvious that

$$\mathsf{P}\left(\sup_{0\le t\le 1}|Y_n(t)-w(t)|>\varepsilon\right) \le \mathsf{P}\left(\max_{0\le k\le n}\left|Y_n\left(\frac{k}{n}\right)-w\left(\frac{k}{n}\right)\right|>\frac{\varepsilon}{2}\right)+$$

$$\mathsf{P}\left(\sup_{0\le t\le 1}\ \sup_{0\le k\le(1/n)}|w(t+k)-w(t)|>\frac{\varepsilon}{2}\right).$$

From this, (12), (14) and (15) we get the statement of the theorem.

Proof of Theorem 2. The proof is similar to the proof of Theorem 1. Let $\{Z_i\}$, $\{X_{k,i}\}$, S_k and $W(t)$ have the same meaning as in the proof of Theorem 1. Now there exists a probability space and a Wiener process $\tilde{W}(t)$ on it, that for every n (see Theorem 1 of [6])

$$\text{(16)}\qquad \mathsf{P}\left(\max_{1\le k\le n}|S_k-\tilde{w}(k)|>x\right) < K_1 e^{-\mu(x-C_1\log n)},$$

where C_1, K_1, μ depend only on the distribution of X. Then by (5), (6) and (13) it follows that

$$\mathsf{P}\left(\max_{0\le k\le n}\left|\frac{\mu(k)-b}{\sqrt{nb}}-w\left(\frac{k}{n}\right)\right|>\lambda\right)\le$$

$$\le\frac{n^3 b}{a^2}+K_1 e^{-\mu(1/2\,\lambda\sqrt{nb}\,-C_2\log(nb+a))}+\varkappa_6\frac{nb+a}{a}e^{-(\lambda^2 nb/12a)}$$

for all positive a. For $a=nb$ we get

$$\mathsf{P}\left\{\max_{0\leqq k\leqq n}\left|\frac{\mu(k)-b}{\sqrt{nb}}-w\left(\frac{k}{n}\right)\right|>\lambda\right\}\leqq\frac{n}{b}+K_1\,e^{-\mu\,(1/2\,\lambda\sqrt{nb}-C_2\log nb)}+\varkappa_9\,e^{-(1/12)\,\lambda^2}.$$

Applying this and (10) we have

$$\mathsf{P}\left(\max_{0\leqq k\leqq n}|\mu(0)+\ldots+\mu(k-1)-kb|>a\right)\leqq$$

$$\leqq\frac{n}{b}+K_1\,e^{-\mu\,(1/4\,(a/n)-C_2\log nb)}+\varkappa_9\,e^{-(1/12)(1/4)(a^2/n^3b)}+\varkappa_4\,e^{-\varkappa_5(a/n\sqrt{nb})}.$$

Then by (5), (13) and (16) it follows that

$$\mathsf{P}\left\{\max_{0\leqq l\leqq n}\left|\frac{\mu(k)-b}{\sqrt{nb}}-w\left(\frac{k}{n}\right)\right|>\lambda\right\}\leqq\frac{n}{b}+K_1\,e^{-\mu(1/4(a/n)-C_2\log nb)}+\varkappa_9\,e^{-(1/12)(1/4)(a^2/n^3b)}+$$

$$+\varkappa_4\,e^{-\varkappa_5\,(a/n\sqrt{nb}}+K_1\,e^{-\mu(1/2\,\lambda\sqrt{nb}-C_2\log(nb+a))}+\varkappa_6\,\frac{nb+a}{a}\,e^{-(\lambda^2\,nb/12\,a)}.$$

Finally it can be shown that for $a=\lambda^2 nb\,\dfrac{1}{36\log\dfrac{b}{n}}$ and $\lambda\leqq 6\left(\log\dfrac{b}{n}\right)^{1/2}$

$$\mathsf{P}\left(\max_{0\leqq k\leqq n}\left|\frac{\mu(k)-b}{\sqrt{nb}}-w\left(\frac{k}{n}\right)\right|>\lambda\right)\leqq\frac{n}{b}+K_1\,e^{-\mu\,(1/4\cdot 36\,(\lambda^2(b/\log b/4))-C_2\log nb)}+$$

$$+\varkappa_9\,e^{-\varkappa_{10}\lambda^4 b/n\log^2(b/n)}+\varkappa_4\,e^{-\varkappa_{11}\,\lambda^2\sqrt{b/n}/\log(b/n)}+K_1\,e^{-\mu\,(1/2\,\lambda\,\sqrt{nb}-C_2\log nb)}+$$

$$+\varkappa_{12}\,\frac{\log\dfrac{b}{n}}{\lambda^2}\left(\frac{n}{b}\right)^3.$$

Hence the theorem is proved.

3. The Levy—Prohorov distance

Let ξ and η be two random elements of the space $\mathsf{D}[0, 1]$, then their Levy—Prohorov distance is defined as

$$\Lambda(\xi,\eta)=\inf\{\varepsilon:\ \text{there exists a probability space and } \xi \text{ and } \eta \text{ on it, such that } \mathsf{P}(\sup_{0\leqq t\leqq 1}|\xi(t)-\eta(t)|>\varepsilon)<\varepsilon\}.$$

The following statement can be easily deduced from Theorems 1 and 2.

THEOREM 3. *Set* $b=n^\delta$ $(\delta>0)$.

(i) *If* $E|x-1|^{2+\gamma}<\infty$, *then*

$$\Lambda(Y_n,W)\leqq(1+\delta)C\max\left\{\frac{\log n}{\sqrt{n}},\ \log n\cdot\frac{1}{n^{(\delta-1/4)}}\cdot\frac{4+2\gamma}{5+2\gamma},\ \frac{1}{n^{(\delta+1/2)(\gamma/3+\gamma)}}\right\},$$

where C depends on the distribution of X and γ.

(ii) *If* Ee^{tX} *exists in a neighbourhood of* $t=0$, *then*

$$\Lambda(Y_n,W)\leqq\max\left\{\frac{\log n}{\sqrt{n}},\ D\frac{\delta^2\log^2 n}{n^{(\delta-1/4)}}\right\},$$

where D *depends on the distribution of X.*

I am grateful to professor G. Tusnády for his helpful suggestions.

REFERENCES

[1] BOROVKOV, K. A., О скорости сходимости ветвящихся процессов к диффузионному, *Теор. вероят. и её примен.*, 1985, 30, 3, 468—477.

[2] CSÖRGŐ, M. and RÉVÉSZ, P., *Strong approximation in probability and statistics*, Akadémiai Kiadó, Budapest, 1981. *MR* **84d**: 60050.

[3] FELLER, W., Diffusion processes in genetics, *Proc. Second Berkeley Symp. Math. Statist. Prob., 1950*, Univ. of California Press, Berkeley and Los Angeles, 1951, 227—246. *MR* **13**—671.

[4] GRIMWALL, A., On the convergence of sequences of branching processes, *Ann. Probability* **2** (1974), 1027—1045. *MR* **50** # 14969.

[5] JIRINA, M., On Feller's branching diffusion processes, *Časopis Pěst. Mat.* **94** (1969), 84—90, 107. *MR* **40** # 939.

[6] KOMLÓS, J., MAJOR, P. and TUSNÁDY, G., An approximation of partial sums of independent RV's and the sample DF.I, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **32** (1975), 111—131. *MR* **51** # 11605b.

[7] KOMLÓS, J., MAJOR, P. and TUSNÁDY, G., An approximation of partial sums of independent RV's, and the sample DF. II, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **34** (1976), 33—58. *MR* **53** # 6697.

[8] LAMPERTI, J., Limiting distributions for branching processes, *Proc. Fifth Berkeley Sympos. Math. Statist. and Probability* (Berkeley, Calif., 1965/66), Vol. II.: Contributions to Probability Theory, Part 2, Univ. California Press, Berkeley, Calif., 1967, 225—241. *MR* **36** # 2231.

[9] LINDVALL, T., Convergence of critical Galton—Watson branching processes, *J. Appl. Probability* **9** (1972), 445—450. *MR* **49** # 9966.

[10] SEVAST'JANOV, B. A., *Весвящиеся процессы* (Branching processes), Москва, Наука. *MR* **49** # 9968.

(Received September 29, 1986)

MTA MATEMATIKAI KUTATÓ INTÉZETE
P.O. BOX 127
H—1364 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 277—287

НЕВЫРОЖДЕННЫЕ ПРАВОАЛЬТЕРНАТИВНЫЕ КОЛЬЦА

Ц. ДАШДОРЖ

В работе [1] В. Г. Скосырский доказал, что невырожденная правоальтернативная алгебра над ассоциативно-коммутативным кольцом с $\frac{1}{2}$ альтернативна.

В настоящей работе эта теорема доказывается без ограниченний на кольцо операторов.

Теорема. *Любое невырожденное правоальтернативное кольцо альтернативно.*

Следствие 1. *Простое правоальтернативное кольцо либо альтернативно, либо является ниль-кольцом.*

Известно, что если R — ассоциативная алгебра над ассоциативно-коммутативным кольцом Φ, то операции $x^2=xx$ $x\circ y=xy+yx$, xyx определяют на Φ-модуле R структуру (квадратичной) йордановой алгебры R^+ (см. [2], [3]).

Если йорданова алгебра J изоморфно вложима в R^+ для некоторой ассоциативной алгебры R, то J называется специальной йордановой алгеброй. Отождествим алгебру J с ее образом в R^+, $J\subseteq R$. Подалгебра $\langle J\rangle$ алгебры R, порожденная множеством J, называется ассоциативной обертывающей алгеброй для J.

В дальнейшем мы будем рассматривать только специальные йордановы алгебры.

Алгебра $J^\sharp=1\cdot\Phi+J$ получается из J внешним присоединением единицы.

Φ-подмодуль $I\leqq J$ называется идеалом йордановой алгебры J, если для любых элементов $a\in T$, $x, y\in J^\sharp$ имеем axa, xax, $axy+yxa\in I$.

Известно (см. [3]), что если I, K — идеалы йордановой алгебры J, то Φ-модуль $U_I(K)$, порожденный множеством $\{aka|a\in I,\ k\in K\}$, также является идеалом в J. В частности, $\bar{I}=U_I(J)$ идеал алгебры J. Обозначим $I^{\langle 0\rangle}=I$, $I^{\langle m+1\rangle}=U_{I^{\langle m\rangle}}(I^{\langle m\rangle})$.

Элемент $a\in J$ называется абсолютным делителем нуля йордановой алгебры J, если $aJ^\sharp a=0$. Йорданова алгебра невырождена, если она не содержит ненулевых абсолютных делителей нуля. Наименьший идеал $M(J)$ алгебры J, фактор-алгебра по которому невырождена, называется радикалом Маккриммона алгебры J.

Йорданова алгебра J первична, если для любых двух идеалов K, L алгебры J из $U_L(K)=0$ следует либо $K=0$ либо $L=0$.

Если J — первичная невырожденная йорданова алгебра, $a\in J$ и I — идеал в J, то из $aIa=0$ следует, что либо $a=0$, либо $I=0$.

Рассмотрим свободную ассоциативную Φ-алгебру $Ass[X]$ от счетного множества порождающих $X=\{x_1, x_2, \ldots\}$ и свободную специальную йорданову алгебру $SJ[X]\subseteq Ass[X]^+$.

В работе [4] построены вполне характеристический идеал T и функция натурального аргумента $f(k)$, $k\geqq 3$, $f(3)=0$, такие, что для любых элементов $t\in T^{\langle f(k)\rangle}$, $a_1, \ldots, a_k\in SJ[X]$ имеем

$$\{a_1 \ldots a_i ta_{i+1} \ldots a_k\} = a_1 \ldots a_i ta_{i+1} \ldots a_k + a_k \ldots a_{i+1} ta_i \ldots a_1 \in SJ[X], \quad 1 \leqq i \leqq k.$$

Обозначим

$$T_m = T^{\langle f(m)\rangle}.$$

Пусть J — первичная невырожденная специальная йорданова алгебра, R — ее ассоциативная обертывающая алгебра. Предположим, что $T(J)\neq 0$.

На алгебре R определим соответствие $*$ следующим образом: если $x\in R$, $x=\sum a_{i1}\ldots a_{ik}$, где $a_{ij}\in J$, то положим $x^*=\sum a_{ik}\ldots a_{i1}$. Ввиду неоднозначности записи элемента x это соответствие не является, вообще говоря, отображением.

Лемма 1. *Пусть $r\in R$ и $rT_m(J)r^*=0$ для некоторого элемента r^* и натурального числа $m\geqq 1$. Тогда существует натуральное число $s\geqq 1$, такое, что $r^*T_s(J)r=0$.*

Доказательство. Пусть $r=x_1+\ldots+x_n$, $x_i=a_{i1}\ldots a_{is_i}$, $a_{ij}\in J$, $s=\max\limits_i s_i$, $t\in T_{2s}(J)$. Тогда

$$r^*tr = (x_1^*+\ldots+x_n^*)\, t(x_1+\ldots+x_n) = \sum_{i\neq j} (x_i^*\, tx_j+x_j^*\, tx_i)+\sum_i x_i^*\, tx_i\in I.$$

Имеем $r^*trT_m(J)r^*tr=0$. Отсюда в силу первичности и невырожденности алгебры J получаем $r^*tr=0$, то есть $r^*T_{2s}(J)\, r=0$. Лемма доказана.

Множество

$$\mathrm{Ann}\,(R, T_m) = \{a\in R| a\, \check{T}_m(J) = \check{T}_m(J)\, a = 0\}$$

является идеалом в R. Действительно, пусть $a\in \mathrm{Ann}\,(R, T_m)$, $x\in J$. Тогда

$$ax\, \check{T}_m(J)\subseteq a\big(x\circ \check{T}_m(J)\big)-a\, \check{T}_m(J)\, x = 0.$$

Положим

$$\mathrm{Ann}(R, T^\infty) = \bigcup_{m\geqq 1} \mathrm{Ann}(R, T_m).$$

Обозначим $H_0=\{r+r^*, rr^*|r\in R\}$,

$$H = \{\Sigma(a_1\ldots a_n + a_n\ldots a_1)+a|a_i, a\in J\}.$$

Лемма 2. *Пусть* $r\in R$, $rT_m(J)r^*=0$ *для некоторых* $m\geqq 1$ *и* r^*. *Тогда* $rHr^*\subseteq \mathrm{Ann}(R, T^\infty)$.

Доказательство. Пусть $h\in H$, $l=rhr^*=\sum_i (x_{i1}\dots x_{ik_i}+x_{ik_i}\dots x_{i1})+y$, $x_{ij}, y\in J$, $k=\max_i k_i$. Для любого элемента $t\in T_{k+1}(J)$ имеем $tlt\in J$. По лемме 1 существует $s\geqq 1$, такое, что $r^*T_s(J)r=0$. Значит, $tltT_s(J)tlt\subseteq tlT_s(J)lt=trhr^*T_s(J)rhr^*t=0$. Теперь $tlt=0$ в силу первичности и невырожденности алгебры J. Для любых элементов $a\in J$, $t'\in T_s(J)$ имеем $(lta+atl)t'(lta+atl)==ltat'lta+atlt'atl=l(tat'+t'at)lta+atl(t'at+tat')l-lt'atlta-atltat'l=0$. Значит $(lta+atl)T_s(J)(lta+atl)=0$. Поскольку $lta+atl\in J$, то $lta+atl=0$.

Далее, $ltat=(lta+atl)t-atlt=0$. Аналогично $tatl=0$. Мы доказали, что $l\check{T}_{k+1}(J)=\check{T}_{k+1}(J)l=0$, то есть $l\in \mathrm{Ann}(R, T_{k+1})\subseteq \mathrm{Ann}(R, T^\infty)$. Лемма доказана.

Следствие 2. *Если* $rT_m(J)r^*=0$, *то*

$$r^*Hr\subseteq \mathrm{Ann}(R, T^\infty).$$

Лемма 3. *Пусть* R — *полупервичная ассоциативная обертывающая алгебра йордановой алгебры* J, $k\in R$ *и* $kHk=0$. *Тогда* $k=0$.

Доказательство. а) Считаем, что R первична. Множество $I=\{a\in R|a^*=0\}$ является идеалом в R. Для $a\in I$ имеем $a=a+a^*\in H$, т.е. $I\subseteq H$, $kIk=0$. Отсюда в силу первичности алгебры R либо $k=0$, либо $I=0$. Если $I=0$, то соответствие $a\to a^*$ является однозначным отображением и, следовательно, инволюцией.

Справедливо следующее утверждение: если R — первичная ассоциативная алгебра с инволюцией $*\colon R\to R$, являющаяся ассоциативной обертывающей алгеброй йордановой алгебры J, $H_0\subseteq J\subseteq H(R, *)=\{a\in R|a^*=a\}$, $k\in R$ и $kJk=0$, то $k=0$. Его доказательство (принадлежащее Е. И. Зельманову) мы приводим в приложении.

б) Пусть R полупервична и $k\neq 0$. Тогда найдется первичный идеал P, такой, что $k\notin P$. При каноническом гомоморфизме $^-\colon R\to \bar{R}=R/P$, $J\to \bar{J}$, $H\to \bar{H}$ получаем $\bar{k}\bar{H}\bar{k}=\bar{0}$. В силу а) $\bar{k}=\bar{0}$, т. е. $k\in P$ что противоречит нашему предположению. Отсюда $k=0$. Лемма доказана.

Лемма 4. *Пусть* J — *первичная невырожденная специальная йорданова алгебра*, R — *ее ассоциативная обертывающая алгебра и* $T(J)\neq 0$. *Пусть* $r\in R$ *и существуют* r^* *и натуральное число* m, *такие, что* $rT_m(J)r^*=0$. *Тогда* $rRr^*\subseteq \bar{\bar{B}}(R)$, где $\bar{\bar{B}}(R)$ — *полный прообраз радикала Бэра* $B(\bar{R})$ *при гомоморфизме*

$$R\to \bar{R}=R/\mathrm{Ann}(R, T^\infty).$$

Доказательство. Из первичности и невырожденности алгебры J следует, что $J\cap \mathrm{Ann}(R, T^I)=0$. Таким образом, $\bar{R}$ является ассоциативной обертывающей алгеброй для J. Кроме того $\bar{r}\bar{H}\bar{r}^*=\bar{r}^*\bar{H}\bar{r}=0$, так как по лемме 2 и следствию 2 rHr^*, $r^*Hr\subseteq \mathrm{Ann}(R, T)$. Пусть $\bar{\bar{R}}=\bar{R}/B(\bar{R})$, где $B(\bar{R})$ — радикал Бэра алгебры $\bar{R}$. Рассмотрим элементы $x\in R$ и $k=rxr^*$, $\bar{k}=\bar{r}\bar{x}\bar{r}^*$, $\bar{\bar{k}}=\bar{\bar{r}}\bar{\bar{x}}\bar{\bar{r}}^*$. Тогда $\bar{k}\bar{H}\bar{k}=0$, так как $\bar{r}^*\bar{H}\bar{r}=0$ и тем более $\bar{\bar{k}}\bar{\bar{H}}\bar{\bar{k}}=0$. В силу полупервичности алгебры $\bar{R}$ по лемме 3 получаем, что $\bar{k}=0$. Значит, $\bar{k}\in B(\bar{R})$, $k\in\bar{B}(R)$. В силу произвольности элемента $x\in R$ получаем $rRr^*\subseteq\bar{B}(R)$. Лемма доказана.

Если A — правоальтернативная алгебра, то операции $x^2=xx$, $xy\cdot x$ определяют на A структуру квадратичной йордановой алгебры A^+. Известно (см. [2]), что алгебра правых умножений $R(A)$ алгебры A на $A^\sharp$ является ассоциативной обертывающей алгеброй для A^+.

Образ элемента $a\in A$ при каноническом вложении в алгебру $R(A)$ обозначим через a'; $(a,b)=a'b'-(ab)'$ и $(a,b)^*=b'a'-(ab)'$ — операторы Смайли.

Лемма Скосырского. *Пусть A — правоальтернативная алгебра, $a,b\in A$. Тогда $(a,b)T(A^+)'(a,b)^*=0$.*

Для полноты изложения приведем ее доказательство, принадлежащее В. Г. Скосырскому.

Элементы из $SJ[X]$ называются йордановыми многочленами. Известно (см. [2]), что для произвольного йорданового многочлена $f(x_1,\dots,x_n)$ и произвольных элементов $a_1,\dots,a_n\in A$ справедливо $f(a_1',\dots,a_n')=(1f(a_1',\dots,a_n'))'$. Если $t(x_1,\dots,x_m)\in T(SJ[X])$; $a,b,a_1,\dots,a_m\in A$, то $(a,b)t'(a_1,\dots,a_m)(a,b)^*$ — йорданов многочлен от $a',b',(ab)',a_1',\dots,a_m'$. Следовательно,

$$(a,b)\,t'(a,b)^* = \big(1(a,b)\,t'(a,b)^*\big)' = 0$$

поскольку $1(a,b)=0$. Лемма доказана.

Из леммы 4 и леммы В. Г. Скосырского вытекает

Следствие 3. *Пусть A — правоальтернативная алгебра, такая, что A^+ — первичная невырожденная йорданова алгебра и $T(A^+)\neq0$; $a,b\in A$. Тогда*

$$(a,b)\,R(A)(a,b)^*,\ (a,b)^*R(A)(a,b)\subseteq\tilde{B}\big(R(A)\big).$$

Лемма 5. *Предположим, что алгебра A удовлетворяет условиям следствия 3. Тогда для любых элементов $a,b,x,y\in A$, операторов $V,W\in R(A)$ справедливо включение*

$$(x,y)\,W\,(a,b)^*\,V(x,y)\in(A,A)\,W(x,y)^*\,V(A,A)+\bar{B}\big(R(A)\big).$$

Доказательство. Для любых операторов $V,W\in R(A)$ пишем $V\equiv W$, если $V-W\in\bar{B}(R(A))$. Полной линеаризацией сравнения $(x,y)W(x,y)^*\equiv0$ получим

$$(x,y)\,W(a,b)^*+(x,b)\,W(a,y)^*+(a,y)\,W(x,b)^*+(a,b)\,W(x,y)^*\equiv0.$$

Используя также частичные линеаризации сравнения $(x,y)^*W(x,y)\equiv0$, получим

$$(x,y)W(a,b)^*\,V(x,y)\equiv-(x,b)W(a,y)^*\,V(x,y)-(a,y)W(x,b)^*\,V(x,y)-$$

$$-(a,b)W(x,y)^*\,V(x,b)\equiv(x,b)W(x,y)^*\,V(a,y)+(a,y)W(x,y)^*\,V(x,b)-$$

$$-(a,b)W(x,y)^*\,V(x,y)\in(A,A)W(x,y)^*\,V(A,A).$$

Лемма доказана.

Предложение 1. *Пусть A — правоальтернативная алгебра, такая, что A^+ — первичная невырожденная йорданова алгебра и $T(A^+)\neq 0$. Тогда A ассоциативна.*

Доказательство. Во всякой алгебре A справедливо тождество Тейхмюллера (см. [2])

$$w(x, y, z)-(wx, y, z)+(w, xy, z)-(w, x, yz)+(w, x, y)z = 0$$

где $x, y, z, w\in A$; $(a, b, c)=(ab)c-a(bc)$ — ассоциатор элементов a, b, c. Обозначим $p=(x, y, z)$. В силу тождества Тейхмюллера

$$\text{(I)} \qquad p' = (x, y, z)' = x'(y, z)-(xy, z)+(x, yz)-(x, y)z'$$

$$\text{(I*)} \qquad p' = (x, y, z)' = (y, z)^* x'-(xy, z)^*+(x, yz)^*-z'(x, y)^*.$$

Пусть $V, W\in R(A)$. Рассмотрим элемент $d=(Vp'W)^{17}$ и запишем вместо p' на нечетном месте его выражение (I), а на четном месте его выражение (I*). После раскрытия скобок в каждом слагаемом хотя бы один из операторов (y, z), (xy, z), (x, yz), (x, y) встретится не меньше 3 раз. Между любыми двумя операторами из (A, A) стоит оператор из $(A, A)^*$. В силу леммы 5 и следствия 3 $d\in \check{B}(R(A))$, откуда следует $p'\in \check{B}(R(A))$.

Покажем, что $p=0$. Предположим, что это не так и построим бесконечную последовательность ненулевых элементов $p_0, p_1, \ldots\in A$ следующим образом. Положим $p_0=p$. Предположим, что элементы $p_0, p_1, \ldots, p_n$ построены. В силу невырожденности алгебры A^+ найдется элемент $x_n\in A$, такой, что $p_n x_n \cdot p_n\neq 0$. Полагаем $p_{n+1}=p_n x_n\cdot p_n$. Последовательность $\overline{p'_0}, \overline{p'_1}, \ldots$ является m-последовательностью в алгебре $\overline{R(A)}=R(A)/\mathrm{Ann}\,(R(A), T^\infty)$. В [5] доказано, что всякая m-последовательность, начинающаяся с элемента из радикала Бэра, конечна. Следовательно, найдется номер $m\geqq 1$, такой, что $p'_m\in \mathrm{Ann}\,(R(A), T^\infty)$. Для некоторого номера $n\geqq 1$ имеем $p'_m(\check{T}_n(A^+))'=0$ и $p_m\check{T}_n(A^+)p_m=0$ в алгебре A^+, что противоречит первичности и невырожденности алгебры A^+. Предложение доказано.

Ниже нам понадобится следующее утверждение, принадлежащее А. Тэди [6].

Лемма Тэди. *Пусть A — правоальтернативная алгебра, I — идеал йордановой алгебры A^+, фактор-алгебра A^+/I по которому невырождена, Тогда I — идеал алгебры A.*

Для ее доказательства А. Тэди заметил, что в произвольной правоальтернативной алгебре выполняется тождество

$$(ar\cdot x)\cdot ar = ar\circ V_{a,x}(r)+xr\circ V_a(r)-V_{a,x}V_r(a)-V_aV_r(x)-V_{xr,ar}(a),$$

где $V_a(x)=ax\cdot a$, $V_{a,b}(x)=ax\cdot b+bx\cdot a$. Из него следует, что если $a\in I$, $r\in A$, то $ar+I$ — абсолютный делитель нуля фактор-алгебры A^+/I. Значит, $ar\in I$. Далее, $ra=(ra+ar)-ar\in I$.

Лемма 6. *Предположим, что правоальтернативная алгебра B удовлетворяет тождеству $xy+yx=0$ и порождена элементами $b_1, \ldots, b_n$, такими, что $b_i^2=0$. Тогда алгебра B нильпотентна.*

Доказательство. В силу $-yx\cdot z=xy\cdot z=-xz\cdot y=y\cdot xz$ алгебра B антиассоциативна. Теперь $B^{n+1}=0$ по лемме из [2]. Лемма доказана.

Лемма 7. *Пусть A — правоальтернативная алгебра с 1 над полем Φ; t: $A\to\Phi$ и n: $A\to\Phi$ — линейная и квадратичная формы соответственно, причем для любого элемента $a\in A$ справедливо равенство $a^2-t(a)a+n(a)\cdot 1=0$. Тогда алгебра A локально конечномерна.*

Доказательство. Пусть $a_1, \ldots, a_n$ — произвольные элементы алгебры A. Наша задача — доказать, что порожденная ими подалгебра $\langle a_1, \ldots, a_n\rangle$ конечномерна. Рассмотрим для этого свободную правоальтернативную алгебру $F=\Phi\langle x_1, \ldots, x_n\rangle$ и ее идеал I, порожденный элементами $vw+wv$, x_i^2, $i=1, \ldots, n$ где v, w — произвольные слова от $\{x_i\}$. По лемме 6 факторалгебра F/I нильпотента степени $\leqq n+1$, т. е. любое слово v длины $m\geqq n+1$ относительно $\{x_i\}$ лежит в I,

$$v=\Sigma(v_i w_i+w_i v_i)V_i+\Sigma x_i^2 W_i,$$

где W_i, V_i — слова от операторов правого и левого умножения на элементы из F.

В алгебре A выполняется равенство $a\circ b-t(a)b-t(b)a+f(a, b)\cdot 1=0$, где

$$f(a, b)=n(a+b)-n(a)-n(b)\in\Phi.$$

Пусть —: $F\to\langle a_1, \ldots, a_n\rangle$ — гомоморфизм, индуцированный отображением $x_i\to a_i$, $1\leqq i\leqq n$. Тогда

$$\bar v=\Sigma(\bar v_i\bar w_i+\bar w_i v_i)\overline{V}_i+\Sigma a_i^2\,\overline{W}_i=$$
$$=\Sigma\big(t(\bar v_i)\bar w_i+t(\bar w_i)\bar v_i-f(\bar v_i, \bar w_i)\cdot 1\big)\overline{V}_i+\Sigma\big(t(a_i)\,a_i-n(a_i)\cdot 1\big)\overline{W}_i.$$

Таким образом, любое слово от $a_1, \ldots, a_n$ длины $\geqq n+1$ линейно выражается через слова длины $<n$ относительно $a_1, \ldots, a_n$. Значит, $\dim_\Phi\langle a_1, \ldots, a_n\rangle<J$. Лемма доказана.

Элемент a йордановой алгебры J называется собственно нильпотентным, если для любого элемента $b\in J$ найдется натуральное число $n(b)$, что $aba\ldots$ $\ldots ba=0$, где a встречается n раз. Алгебра называется собственно нильалгеброй, если всякий ее элемент собственно нильпотентен. Максимальный собственно ниль-идеал йордановой алгебры J называется ее собственным ниль-радикалом и обозначается через $N(J)$. Если A — правоальтернатиная алгебра, то в силу леммы Тэди $N(A)=N(A^+)$ — идеал алгебры A.

Лемма 8. *Предположим, что правоальтернативная Φ-алгебра A локально конечномерна и $N(A)=0$. Тогда алгебра A альтернативна.*

Доказательство. Пусть $\mathrm{alt}(A)$ — идеал алгебры A, порожденный элементами вида (a, a, b), $a, b\in A$. Покажем, что $\mathrm{alt}(A)\subseteq N(A)$. Пусть $x\in\mathrm{alt}(A)$, $y\in A$. Тогда $x=\sum(a_i, a_i, b_i)W_i$, где $a_i, b_i\in A$; W_i — слова от операторов правых и левых умножений на элементы из A. Рассмотрим подалгебру A_1 алгебры A, порожденную элементами a_i, b_i, y и всеми элементами, входящими в запись операторов W_i. По теореме Тэди [7] фактор-алгебра $A_1/N(A_1)$ альтернативна. Значит, $x \in N(A_1)$ и найдется натуральное число n, такое, что $xyx\ldots$ $\ldots yx=0$, где x встречается n раз. Мы доказали, что элемент x собственно нильпотентен в алгебре A^+, то есть $\operatorname{alt}(A) \subseteq N(A)=0$. Лемма доказана.

Йорданова алгебра называется примитивной, если она содержит максимальный модулярный внутренний идеал (см. [3] не), содержащий ненулевых идеалов этой алгебры.

Правоальтернативная алгебра A называется примитивной, если примитивна йорданова алгебра A^+.

Пусть $\operatorname{Jac}(A^+)$ и $M(A^+)$ — радикалы Джекобсона и Маккриммона алгебры A^+ соответственно. По лемме Тэди $\operatorname{Jac}(A^+)$ и $M(A^+)$ — идеалы алгебры A.

Предположим, что $\operatorname{Jac}(A^+)=0$. Полупростая йорданова алгебра A^+ аппроксимируется примитивными йордановыми алгебрами (см. [3]), т. е. найдутся семейство идеалов $\{P_\alpha \lhd A^+ | \alpha \in M\}$, такое, что для любого $\alpha \in M$ йорданова алгебра A^+/P_α примитивна и $\bigcap_{\alpha \in M} P_\alpha=0$.

Поскольку всякая примитивная йорданова алгебра первична и невырождена, то по лемме Тэди P_α — идеал алгебры A. Таким образом, полупростая правоальтернативная алгебра аппроксимируется примитивными правоальтернативными алгебрами.

Предложение 2. *Примитивная правоальтернативная алгебра A над алгебраически замкнутым полем Φ,* $\operatorname{card} \Phi > \dim_\Phi A$ *альтернативна.*

Доказательство. Если $T(A^+) \neq 0$, то в силу предложения 1 алгебра A ассоциативна. Если $T(A^+)=0$, то из результатов работы [4] следует, что A^+ — йорданова алгебра симметрической билинейной формы в векторном пространстве над некоторым полем $\Gamma \supseteq \Phi$. В силу ограничений на поле Φ, $\Gamma=\Phi$. Следовательно, определены линейная форма t: $A \to \Phi$ и квадратичная форма n: $A \to \Phi$ такие, что $a_2-t(a)a+n(a)\cdot 1=0$ для любого элемента $a \in A$. По лемме 7 алгебра A локально конечномерна. Поскольку примитивная йорданова алгебра собственно ниль-полупроста, то по лемме 8 алгебра A альтернативна. Предложение доказано.

Лемма 9. *Пусть Φ — произвольное ассоциативно-коммутативное кольцо, F — свободная правоальтернативная Φ-алгебра от бесконечного множества порождающих $\{x_i | i \in I\}$. Тогда $N(F)=M(F)$.*

Доказательство. Включение $M(J) \subseteq N(J)$ справедливо для любой йордановой алгебры J (см. [2]). Покажем, что $N(F) \subseteq M(F)$. Пусть $a=a(x_1, \ldots$ $\ldots, x_n) \in N(F)$ и порождающий элемент x_{n+1} не входит в запись элемента a. В силу собственной нильпотентности элемента a найдется натуральное число m, такое, что $ax_{n+1}a\ldots x_{n+1}a=0$ где элемент a встречается m раз. Тогда для любого элемента $b \in F$ выполняется равенство $aba\ldots ba=0$. Как доказано в [8], отсюда следует, что $a \in M(F^+)$. Лемма доказана.

Предложение 3. *Невырожденная правоальтернативная алгебра над полем Φ альтернативна.*

Доказательство. Достаточно доказать включение $\mathrm{alt}(F) \subseteq M(F)$ для свободной правоальтернативной Φ алгебры F от бесконечного множества порождающих X. Пусть $\bar{\Phi}$ — алгебраически замкнутое расширение поля Φ, $\mathrm{card}\,\bar{\Phi} > \mathrm{card}\,X$, $\bar{F} = F \otimes_{\Phi} \bar{\Phi}$. Полупростая алгебра $\bar{F}/\mathrm{Jac}(\bar{F})$ аппроксимируется примитивными правоальтернативными алгебрами. В силу предложения 2 алгебра $\bar{F}/\mathrm{Jac}(\bar{F})$ альтернативна.

Используя кардинальный трюк Амицура (см. [9], [3]), получаем $\mathrm{Jac}(\bar{F}) = N(\bar{F})$. Теперь $\mathrm{alt}(F) \subseteq \mathrm{alt}(\bar{F}) \cap F \subseteq \mathrm{Jac}(\bar{F}) \cap F = N(\bar{F}) \cap F \subseteq N(F) = M(F)$. Предложение доказано.

Напомним, что правоальтернативное кольцо называется первичным, если для любых его идеалов K, L из $KL=0$ следует, что либо $K=0$, либо $L=0$.

Лемма 10. *Первичное невырожденное правоальтернативное кольцо A альтернативно.*

Доказательство. Из первичности кольца A следует, что либо

1) A не имеет аддитивного кручения, либо

2) найдется простое число $p \geqq 2$, такое, что $pA=0$.

В первом случае кольцо дробей $\tilde{A} = \{a/n \mid a \in A,\ n \in \mathbf{Z},\ n \neq 0\}$ есть невырожденная правоальтернативная алгебра над полем рациональных чисел $\mathbf{Q}$. Во втором случае A — невырожденная правоальтернативная алгебра над полем вычетов $\mathbf{Z}(p)$. Теперь наше утверждение следует из предложения 3. Лемма доказана.

Для произвольного правоальтернативного кольца A обозначим через $\mathrm{Мих}(A)$ его локально r_1-нильпотентный радикал (радикал Михеева, см. [2]). Из теорем А. М. Слинько и В. Г. Скосырского (см. [2]) следует, что $M(A) \subseteq \mathrm{Мих}(A)$.

Доказательство теоремы. Пусть F — свободное правоальтернативное кольцо от бесконечного множества порождающих. Так как $\mathrm{Мих}(F)$ — нильрадикал, то по лемме 9 $M(F) = \mathrm{Мих}(F) = N(F)$. По теореме И. М. Михеева [10] фактор-кольцо $F/\mathrm{Мих}(F)$ аппроксимируется первичными Мих-полупростыми кольцами. Поскольку всякое первичное Мих-полупростое кольцо невырождено, то в силу леммы 10 оно альтернативно. Значит, альтернативно и кольцо $F/\mathrm{Мих}(F)$, $\mathrm{alt}(F) \subseteq M(F)$. Теорема доказана.

Я благодарю Е. И. Зельманова, К. Маккриммона, В. Г. Скосырского, познакомивших меня с рукописями своих работ до их опубликования. Пользуюсь случаем поблагодарить также Л. А. Бокутя за руководство работой и Е. И. Зельманова за ценные советы.

ПРИЛОЖЕНИЕ

Предложение (Е. И. Зельманов). *Пусть R — полупервичная ассоциативная алгебра над полем Φ с инволюцией $*: R \to R$, $a \in R$ и для произвольного элемента $r \in R$ выполняется $a(r+r^*)a = arr^*a = 0$. Тогда множество $H_0 = \{r+r^*, rr^* / r \in R\}$ централизует идеал $Ug_R(a)$, порожденный элементом a в алгебре R.*

Приведем для полноты изложения доказательство этого результата, также принадлежащее Е. И. Зельманову.

Пусть A — ассоциативная алгебра, $t\in A$. Зададим на Φ — пространстве A новое умножение, полагая $x*y=xty$ и обозначим новую алгебру через $A^{(t)}$. Множество $\mathrm{Ker}\,A^{(t)}=\{x\in A\,|\,txt=0\}$ — идеал алгебры $A^{(t)}$.

Лемма 1. *Если алгебра A примитивна, то алгебра $A^{(t)}/\mathrm{Ker}\,A^{(t)}$ также примитивна.*

Доказательство. Пусть V_A — точный неприводимый правый A-модуль. Зададим на подпространстве Vt структуру $A^{(t)}$-модуля, полагая

$$(Vt, A^{(t)})\ni(v, x) \to vxt\in Vt.$$

Легко видеть, что $Vt\,\mathrm{Ker}\,A^{(t)}=0$ и Vt — точный $A^{(t)}/\mathrm{Ker}\,A^{(t)}$-модуль. Покажем, что этот модуль неприводим. Пусть $Vt\ni vt\neq 0$, $v\in V$. Тогда в силу неприводимости A-модуля V имеем $vtA=V$ и, следовательно, $vtAt=Vt$. Значит, $vtA^{(t)}=Vt$. Лемма доказана.

Лемма 2. *Предположим, что в условиях предложения 1 поле Φ алгебраи чески замкнуто и Φ-алгебра R конечномерна. Тогда $[H_0, Ug_R(a)]=0$.*

Доказательство. Не уменьшая общности, можно считать, что R есть либо простая алгебра Φ_n, либо прямая сумма простых антиизоморфных алгебр $\Phi_m\oplus\Phi_m^0$ с обменной инволюцией $*\colon a\oplus b^0\to b\oplus a^0$. Во втором случае из равенства $aH_0a=0$ следует $a=0$. В первом случае, если $n\neq 2$ или инволюция $*\colon \Phi_2\to\Phi_2$ не является симплектической, то равенство $aHa=0$ снова влечет $a=0$ (см. [3]). Если $*\colon \Phi_2\to\Phi_2$ — симплектическая инволюция, то множество H_0 состоит из скалярных матриц. Лемма доказана.

Лемма 3. *Предположим, что в условиях предложения 1 поле Φ алгебраически замкнуто и Φ-алгебра R локально конечномерна. Тогда $[H_0, Ug_R(a)]\subseteq$ $\subseteq N(R)$.*

Доказательство. Для произвольных элементов $a_1, \ldots, a_n\in R$ и сопряженных с ними элементов $a_1^*, \ldots, a_n^*$ рассмотрим порожденную ими конечномерную подалгебру $R_1=\langle a_1, \ldots, a_n, a_1^*, \ldots, a_n^*, a\rangle$. По лемме 2 $[H_0(R_1, *), Ug_{R_1}(a)]\subseteq N(R_1)$. В силу произвольности элементов $a_1, \ldots, a_n\in R$ отсюда следует утверждение леммы.

Лемма 4. $[H_0, Ug_R(a)]\subseteq N(R)$.

Доказательство. Пусть F — алгебраически замкнутое расширеиие поля Φ большой мощности $\mathrm{card}\,F>\dim_\Phi R$. Переходя к алгебре $R\otimes_\Phi F/N(R\otimes_\Phi F)$ можно не уменьшая общности считать, что поле Φ — алгебраически замкнуто и $\mathrm{card}\,\Phi>\dim_\Phi R$. В этом случае по теореме Амицура [9] $\mathrm{Jac}\,(R)=$ $=N(R)$.

Покажем, что для произвольного элемента $r\in R$ выполняется равенство $ara^*H_0ara^*=0$. В самом деле, если $h\in H_0$, то

$$ara^*\,hara^* = a(ra^*\,h+har^*)\,ara^*\,ahar^*\,ara^* = 0.$$

В частности, для любого $h\in H_0$ элемент $a_1=aha^*$ есть абсолютный делитель нуля йордановой алгебры H_0. Для произвольного элемента $r\in R$ имеем $a_1ra_1ra_1=a_1(r+r^*)a_1ra_1-a_1r^*a_1ra_1=0$, поскольку $r^*H_0r\subseteq H_0$. По теореме Левицкого $a_1R\subseteq B(R)=0$. Мы доказали, что алгебра H_0 невырождена. Значит $aH_0a^*=0$ и $(a+a^*)H_0(a+a^*)=0$. Снова ввиду невырожденности алгебры H_0 имеем $a^*=-a$.

Для произвольных элементов $x, y\in R$ имеем

$$\begin{aligned} axaya &= a(xay+y^*a^*x^*)a-ay^*a^*x^*a = \\ &= ay^*ax^*a = a(y^*+y)ax^*a-ayax^*a = \\ &= -aya(x+x^*)a+ayaxa = ayaxa. \end{aligned}$$

Пусть P — примитивный идеал алгебры R, $\bar{\ }: R\to R/P$ — естественный гомоморфизм. По лемме 1 $\bar{R}^{(a)}/\mathrm{Ker}\,\bar{R}^{(\bar{a})}$ — примитивная коммутативная Φ-алгебра. Следовательно, $\bar{a}\bar{R}\bar{a}\subseteq\Phi\bar{a}$. Рассмотрим гомоморфизм

$$=: R\to R/P\cap P^*.$$

В силу сказанного выше $\dim_\Phi \bar{\bar{a}}\bar{\bar{R}}\bar{\bar{a}}\leqq 2$. Элемент $\bar{\bar{a}}$ порождает в алгебре $\bar{\bar{R}}$ локально конечный идеал $\bar{\bar{I}}$. По лемме 3 $[Ug_{\bar{\bar{I}}}(\bar{\bar{a}}), \bar{\bar{I}}\cap\bar{\bar{H}}]=0$. Очевидно, $\bar{\bar{I}}^3\subseteq Ug_{\bar{\bar{I}}}(\bar{\bar{a}})$, поэтому $[\bar{\bar{I}}^3, H_0(\bar{\bar{I}}, *)]=0$. Ввиду полупервичности алгебры $\bar{\bar{R}}$ отсюда следует $[\bar{\bar{I}}, H_0(\bar{\bar{I}}, *)]=0$. Далее. $\bar{\bar{I}}[\bar{\bar{R}}, H_0(\bar{\bar{I}}, *)]\subseteq[\bar{\bar{I}}, H_0(\bar{\bar{I}}, *)]+[\bar{\bar{I}}, H_0(\bar{\bar{I}}, *)]\bar{\bar{R}}=0$. Значит, $[\bar{\bar{R}}, H_0(\bar{\bar{I}}, *)]=0$.

Пусть $\bar{\bar{h}}\in H_0(\bar{\bar{R}}, *)$, $\bar{\bar{h}}_1\in H_0(\bar{\bar{I}}, *)$, $\bar{\bar{x}}\in\bar{\bar{R}}$. В силу центральности элемента $\bar{\bar{h}}_1$ произведение $\bar{\bar{h}}\bar{\bar{h}}_1^2$ лежит в $H_0(\bar{\bar{I}}, *)$. В самом деле, если $\bar{\bar{h}}=\bar{\bar{r}}+\bar{\bar{r}}^*$, то $\bar{\bar{h}}\bar{\bar{h}}_1^2=(\bar{\bar{r}}\bar{\bar{h}}_1^2)+(\bar{\bar{r}}\bar{\bar{h}}_1^2)^*$. Если же $\bar{\bar{h}}=\bar{\bar{r}}\bar{\bar{r}}^*$, то $\bar{\bar{h}}\bar{\bar{h}}_1^2=(\bar{\bar{r}}\bar{\bar{h}}_1)(\bar{\bar{r}}\bar{\bar{h}}_1)^*$. Поэтому $[\bar{\bar{x}}, \bar{\bar{h}}]\bar{\bar{h}}_1^2=[\bar{\bar{x}}, \bar{\bar{h}}\bar{\bar{h}}_1^2]=0$. Так как элемент $\bar{\bar{h}}_1$ лежит в центре алгебры $\bar{\bar{R}}$, а алгебра $\bar{R}\cong\bar{\bar{R}}/\bar{\bar{P}}$ первична, то либо $[\bar{\bar{R}}, H_0(\bar{\bar{R}}, *)]\subseteq\bar{\bar{P}}$, либо $H_0(\bar{\bar{I}}, *)\subseteq\bar{\bar{P}}$. Если $H_0(\bar{\bar{I}}, *)\subseteq\bar{\bar{P}}$, то для любого элемента $t\in\bar{\bar{I}}$ имеем $t^2=t(t+t^*)-tt^*\in\bar{\bar{P}}$, поэтому $\bar{\bar{I}}\subseteq\bar{\bar{P}}$. В любом случае $[\bar{\bar{I}}, H_0(\bar{\bar{R}}, *)]\subseteq\bar{\bar{P}}$, т.е. $[Ug_R(a), H_0(R, *)]\subseteq P$. Лемма доказана.

Доказательство предложения 1. Пусть $b\in Ug_R(a)$, $h\in H_0(R, *)$, $c=[b, h]$. Наша задача — показать, что $c=0$.

Рассмотрим в множестве натуральных чисел $\mathbf{N}$ ультрафильтр Фреше J и ультрастепень $\tilde{R}=R^N/J$ алгебры R (см. [II]). В алгебре $\tilde{R}$ действиет инволюция $\tilde{*}: (x_i|i\in\mathbf{N})/J\to(x_i^*|i\in\mathbf{N})/J$, причем $H_0(\tilde{R}, \tilde{*})=H_0(R, *)^N/J$. Пусть $\tilde{a}$, $\tilde{b}$, $\tilde{h}$, $\tilde{c}$ — ультрастепени элементов a, b, h, c соответственно. Тогда $\tilde{a}H_0(\tilde{R}, \tilde{*})\tilde{a}=0$, $\tilde{b}\in Ug_{\tilde{R}}(\tilde{a})$ и, значит, $c\in\mathbf{N}(\tilde{R})$. Отсюда следует, что правый идеал cR-ниль-алгебра ограниченной степени. По теореме Левицкого (см. [5]) $c\in B(R)=0$. Предложение доказано.

Следствие 1. *Пусть R — первичное кольцо с инволюцией $*: R\to R$. Если кольцо R содержит ненулевой элемент a, такой, что $aH_0(R, *)a=0$, то его центр $Z(R)$ отличен от нуля и центральное замыкание $Z(R)^{-1}R$ квадратично над полем $Z(R)^{-1}Z(R)$.*

Доказательство. Пусть $a\neq 0$, $aH_0(R, *)a=0$. В силу предложения 1 $[Ug_R(a), H_0(R, *)]=0$. Так как кольцо R первично, то отсюда следует

$[R, H_0(R, *)]=0$, $H_0(R, *)\subseteq Z(R)$. Теперь для любого элемента $x\in R$ имеем $x^2-z_1x+z_2=0$, где $z_1=x+x^*$, $z_2=x^*x\in Z(R)$.

Следствие 2. *Предположим, что первичная алгебра R с инволюцией $*: R\to R$ является ассоциативной обертывающей алгеброй йордановой подалгебры $H_0(R, *)\subseteq J\subseteq H(R, *)$. Если $a\in R$ и для любого элемента $x\in J$ имеем $axa=0$, то $a=0$.*

Доказательство. Ввиду следствия 1 можно предполагать, что R-алгебра 2×2-матриц над полем F,

$$\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}^* = \begin{pmatrix}\delta & -\beta\\ -\gamma & \alpha\end{pmatrix}.$$

Алгебра R порождена множеством $H(R, *)$ лишь в случае $\operatorname{char} F=2$. Теперь заметим, что единственной йордановой подалгеброй алгебры $H(R, *)$, порож дающей R, является сама алгебра $H(R, *)$. Из равенства $aH(R, *)a=0$ следует $a=0$.

ЛИТЕРАТУРА

[1] Скосырский, В. Г., Правоальтернативные алгебры, *Алгебра и логика* **23** (1984), №2, 185—192.

[2] Жевлаков, К. А., Слинько, А. М., Шестаков, И. П. и Ширшов, А. И., *Кольца, близкие к ассоциативным* (Rings that are nearly associative), *Современная алгебра*, Наука, Москва, 1978. *MR* **80h**: 17002.

[3] Jacobson, N., *Structure and representations of Jordan algebras*, Amer. Math. Soc. Colloq. Publ., Vol. **39**, American Mathematical Society, Providence, R. I., 1968. *MR* **40** # 4330.

[4] McCrimmon, K. and Zelmanov, E. I., The structure of prime quadratic Jordan algebras, Munich, 1984 (preprint).

[5] Джекобсон, Н., *Строение колец*, Изд-во иностр. лит., Москва, 1961. *MR* **22** # 11005.

[6] Thedy, A., Radicals of right alternative ar i Jordan rings, *Comm. Algebra* **12** (1984), 857—887. *MR* **85i**: 17030.

[7] Thedy, A., Right alternative rings, *J. Algebra* **37** (1975), 1—43. *MR* **52** # 5758.

[8] Зельманов, Е. И., О первичных йордановых тройных системах, *Сибирск. Mat. Ж.* **24** (1983), 23—37. *MR* **85c**: 17013.

[9] Amitsur, S. A., Algebras over infinite fields, *Froc. Amer. Math. Soc.* **7** (1956), 35—48. *MR* **17**—822.

[10] Михеев, И. М., Локально правонильпотентный радикал в классе правоальтернативных колец, *Ал ебра и логика* **11** (1972), 174—175. *MR* **47** # 1900.

(Поступило 10-ого октября 1986 г.)

КАФЕДРА МАТЕМАТИКИ
МОНГОЛСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
П/Я 46/534
УЛААН БААТАР
МОНГОЛИЯ

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 289—293

ON THE COMPLEMENTEDNESS OF THE LATTICE OF WEAK CONGRUENCES

BRANIMIR ŠEŠELJA and GRADIMIR VOJVODIĆ

In this paper we give necessary and sufficient conditions under which the lattice of all weak congruences (i.e. of all congruences on all subalgebras) of an algebra satisfying CEP and CIP (see below) is complemented.

We give a complete answer to the same question if this complemented lattice is modular, and we also characterize algebras with Boolean lattice of weak congruences.

Let $\mathscr{A}=(A, F)$ be an algebra and $K\subseteq A$ the set of its constants. Then ([1]) a *weak congruence relation* ϱ on $\mathscr{A}$ is a symmetric, transitive and compatible relation on $\mathscr{A}$, satisfying a weak reflexivity: If $c\in K$, then $c\varrho c$.

We shall use the following notations. For a given algebra $\mathscr{A}$,

$S(\mathscr{A})$ is the lattice of all its subalgebras,

$C(\mathscr{A})$ is the lattice of all congruences on $\mathscr{A}$,

$C_w(\mathscr{A})$ is the lattice of all weak congruences on $\mathscr{A}$.

It is clear that the set of all weak congruences on $\mathscr{A}$ coincides with the set of all congruences on all subalgebras of $\mathscr{A}$, i.e.,

$$C_w(\mathscr{A}) = \bigcup\big(C(\mathscr{B})|\mathscr{B}\in S(\mathscr{A})\big).$$

It is also clear that $C_w(\mathscr{A})$ is a lattice under set inclusion.

$\mathscr{A}$ is said to have the *congruence intersection property* (*CIP*) ([1]) if, for all $\varrho, \theta\in C_w(\mathscr{A})$

$$(\varrho\wedge\theta)_A = \varrho_A\wedge\theta_A$$

where

$$\varrho_A \overset{\text{def}}{=} \bigcap\big(\sigma\in C(\mathscr{A})|\varrho \leqq \sigma\big).$$

If d_ϱ is the diagonal of $\varrho\in C_w(\mathscr{A})$ (i.e., $d_\varrho=\varrho\wedge\Delta$ where $\Delta=\{(x, x)|x\in A\}$) then $\varrho_A=\varrho\vee\Delta$, and thus CIP expresses a distributivity property for Δ.

Recall that an algebra $\mathscr{A}$ satisfies the *congruence extension property* (*CEP*) if every congruence on an arbitrary subalgebra of $\mathscr{A}$ is the restriction of some congruence on $\mathscr{A}$.

1980 *Mathematics Subject Classification* (1985 Revision). Primary 06C15; Secondary 08B10.
Key words and phrases. Weak congruence lattice, complementedness.

It was proved in [1] that:

(I) $C_w(\mathscr{A})$ is an algebraic lattice which includes $C(\mathscr{A})$ and $S(\mathscr{A})$ as sublattices (as a filter and an ideal, respectively), the latter up to isomorphism, and that the lattice $S(\mathscr{A})$ is also a homomorphic image of $C_w(\mathscr{A})$.

(By virtue of the above mentioned embedding, the subalgebras are represented in $C_w(\mathscr{A})$ by diagonal relations).

(II) $C_w(\mathscr{A})$ is a modular lattice iff
(i) $S(\mathscr{A})$ is a modular lattice;
(ii) $C(\mathscr{A})$ is a modular lattice;
(iii) A satisfies the CIP;
(iv) A satisfies the CEP.

LEMMA 1. *Let $\mathscr{A}=(A, F)$ be a nontrivial algebra* ($|A|>1$) *such that $C_w(\mathscr{A})$ is a complemented lattice. Then the least subalgebra $\mathscr{B}_m$ of $\mathscr{A}$ is nonvoid.*

REMARK. If $|A|=1$ then either $C_w(\mathscr{A})$ is trivial $(K=A)$ or it is a two-element chain $(K=\emptyset\subset A)$. In the following we consider nontrivial algebras only.

PROOF. Let D' be a complement of Δ. Then $\Delta\vee D'=A^2$ and $\Delta\wedge D'=d_m$, where $d_m=\Delta\wedge B_m^2$.
Since $\Delta\neq A^2$, D' cannot be the empty relation, hence also $d_{D'}\neq\emptyset$. On the other hand, we have by definition $d_{D'}=\Delta\wedge D'$. Thus, $d_m=d_{D'}\neq\emptyset$, which proves that $B_m\neq\emptyset$. □

COROLLARY 2. *If, for an algebra $\mathscr{A}$, $C_w(\mathscr{A})$ is complemented then B_m^2 is a complement of Δ, and if the congruences of $\mathscr{B}_m$ can be extended to $\mathscr{A}$, then this complement of Δ is unique.*

PROOF. Denote by D' a complement of Δ. By the proof of Lemma 1, we have $d_{D'}=d_m$ hence $D'\leq B_m^2$, thus

$$\Delta\vee B_m^2 \geq \Delta\vee D' = A^2 \quad\text{and}\quad \Delta\wedge B_m^2 = d_m,$$

which proves the first claim. Finally, if $D'=\bar{D}\wedge B_m^2$ for a $\bar{D}\in C(\mathscr{A})$, then we have $\bar{D}=\Delta\vee D'\geq\Delta\vee D'=A^2$ and thus

$$D' = \bar{D}\wedge B_m^2 = A^2\wedge B_m^2 = B_m^2. \quad \square$$

REMARK. The preceding statement shows that no nontrivial algebra having a one-element subalgebra (a group, for example) can have a complemented lattice of weak congruences.

PROPOSITION 3. *Let $\mathscr{A}$ be an algebra satisfying CIP and CEP and such that $C_w(\mathscr{A})$ is complemented. Then:*
1) *For every $\mathscr{B}\in S(\mathscr{A})$, the mapping $\varrho\mapsto\varrho_A$ is a lattice isomorphism from $C(\mathscr{B})$ to $C(\mathscr{A})$;*
2) *$C(\mathscr{A})$ is a complemented lattice;*
3) *$S(\mathscr{A})$ is a complemented lattice.*

PROOF. 1) For an arbitrary $\varrho\in C(\mathscr{B})$, by CEP there is a $\bar\varrho\in C(\mathscr{A})$ such that $\varrho=\bar\varrho\wedge B^2$. Now we have:

$$\varrho_A\wedge B^2 \geqq \varrho = \bar\varrho\wedge B^2 \geqq (\bar\varrho\wedge\varrho_A)\wedge B^2 = \varrho_A\wedge B^2,$$

hence $\varrho=\varrho_A\wedge B^2$. Therefore the mapping $\varrho\mapsto\varrho_A$ is an injection from $C(\mathscr{B})$ to $C(\mathscr{A})$. Moreover, the image of $C(\mathscr{B})$ in $C(\mathscr{A})$ is an ideal. Indeed, if $\varrho\in C(\mathscr{B})$, $\theta\in C(\mathscr{A})$, $\theta<\varrho_A$, then it follows that $\theta\wedge\varrho\in C(\mathscr{B})$, and by CIP $(\theta\wedge\varrho)_A=\theta_A\wedge\varrho_A= =\theta\wedge\varrho_A=\theta$. Thus, every $\theta<\varrho_A$ is the image of a congruence in $C(\mathscr{B})$. Furthermore, by Corollary 2, we have

$$(B^2)_A = B^2\vee\Delta \geqq B_m^2\vee\Delta = A^2,$$

and since the image of $C(\mathscr{B})$ is an ideal of $C(\mathscr{A})$, it must be the whole $C(\mathscr{A})$. Finally, $\varrho\mapsto\varrho_A$ is obviously order preserving, which completes the proof of Claim 1.

2) By 1), every congruence in $C(\mathscr{A})$ has the form ϱ_A for some $\varrho\in C(\mathscr{B}_m)$.
Consider a complement ϱ' of ϱ. Then $A^2=\varrho\vee\varrho'$, which implies that $\Delta=d_\varrho\vee d_{\varrho'}$. Since $d_\varrho=d_m$, it follows that $d_{\varrho'}=\Delta$, i.e., $\varrho'\in C(\mathscr{A})$. Now we have

$$\varrho_A\vee\varrho' \geqq \varrho\vee\varrho' = A^2$$

and by CIP,

$$\varrho_A\wedge\varrho'_A = (\varrho\wedge\varrho')_A = (d_m)_A = \Delta.$$

Hence ϱ' is a complement of ϱ_A in $C(\mathscr{A})$.

3) This follows immediately from the fact that $S(\mathscr{A})$ is a homomorphic image of $C_w(\mathscr{A})$ and the latter is complemented. □

The converse of the last proposition is also valid:

PROPOSITION 4. *Let $\mathscr{A}$ be an algebra satisfying conditions* 1), 2) *and* 3) *from Proposition 3. Then $C_w(\mathscr{A})$ is a complemented lattice.*

PROOF. If $\varrho\in C_w(\mathscr{A})$ then there is a (unique) $\mathscr{B}\in S(\mathscr{A})$ such that $\varrho\in C(\mathscr{B})$. Let $\mathscr{C}\in S(\mathscr{A})$ be a complement of $\mathscr{B}$ in $S(\mathscr{A})$ (by 3). Since the lattice $S(\mathscr{A})$ is isomorphic with that of the diagonals, we have $d_\varrho\wedge d_\sigma=d_m$ and $d_\varrho\vee d_\sigma=\Delta$ for any $\sigma\in C(\mathscr{C})$.
Now, let θ be a complement of ϱ in $C(\mathscr{B})$ (by 2)). By 1) we can choose a $\varrho'\in C(\mathscr{C})$ so that $\varrho'_A=\theta_A$.
We shall show that ϱ' is a complement of ϱ in $C_w(\mathscr{A})$.
Indeed, we have $\varrho'\wedge\varrho\in C(\mathscr{B}_m)$ since $d_\varrho\wedge d_{\varrho'}=d_m$. Furthermore,

$$(\varrho\wedge\varrho')_A \leqq \varrho_A\wedge\varrho'_A = \varrho_A\wedge\theta_A \overset{\text{by 1)}}{=} (\varrho\wedge\theta)_A = \Delta.$$

Hence, again by 1), $\varrho\wedge\varrho'=d_m$.
For the supremum, we have $\varrho\vee\varrho'\in C(\mathscr{A})$ since $d_\varrho\vee d_{\varrho'}=\Delta$, and then

$$\varrho\vee\varrho' = \varrho\vee\varrho'\vee\Delta = (\varrho\vee\Delta)\vee(\varrho'\vee\Delta) = \varrho_A\vee\varrho'_A =$$

$$= \varrho_A\vee\theta_A \overset{\text{by 1)}}{=} (\varrho\vee\theta)_A = (B^2)_A = A^2. \quad \square$$

The algebras satisfying condition 1) from Proposition 3 are in the class of those that have CEP and CIP;

PROPOSITION 5. *Let $\mathscr{A}$ be an algebra satisfying* 1) *from Proposition 3. Then $\mathscr{A}$ has CEP and CIP.*

PROOF. 1)$\Rightarrow$CEP: Let $\mathscr{B}\in S(\mathscr{A})$ and $\varrho\in C(\mathscr{B})$, then we have $\varrho\leqq\varrho_A\wedge B^2$ and

$$(\varrho_A\wedge B^2)_A \leqq \varrho_A\wedge B^2_A = \varrho_A\wedge A^2 = \varrho_A,$$

hence

$$\varrho_A\wedge B^2 \leqq \varrho,$$

and thus $\varrho=\varrho_A\wedge B^2$, which proves that $\mathscr{A}$ has CEP.

1)$\Rightarrow$CIP: Take any $\varrho,\theta\in C_w(\mathscr{A})$, say, $\varrho\in C(\mathscr{B})$ and $\theta\in C(\mathscr{C})$, and put $\mathscr{D}=\mathscr{B}\wedge\mathscr{C}$ in $S(\mathscr{A})$. Then there are unique $\varrho',\theta'\in C(\mathscr{D})$ such that $\theta'_A=\theta_A$ and $\varrho'_A=\varrho_A$. We are going to prove that $\varrho'\wedge\theta'=\varrho\wedge\theta$. Indeed, we have $\varrho,\varrho'_B\in C(\mathscr{B})$,

$$\varrho_A = \varrho'_A = \varrho'\vee\Delta = \varrho'\vee d_{B^2}\vee\Delta = (\varrho'_B)_A,$$

hence $\varrho=\varrho'_B$. Therefore $\varrho\geqq\varrho'$, and similarly $\theta>\theta'$; thus $\varrho\wedge\theta\geqq\varrho'\wedge\theta'$. On the other hand,

$$((\varrho\wedge\theta)_B)_A = (\varrho\wedge\theta)_A \leqq \theta_A = \theta'_A = (\theta'_B)_A,$$

hence $(\varrho\wedge\theta)_B\leqq\theta'_B$, and therefore

$$(\varrho\wedge\theta)_B \leqq \varrho\wedge\theta'_B = \varrho'_B\wedge\theta'_B = (\varrho'\wedge\theta')_B,$$

whence $\varrho\wedge\theta\leqq\varrho'\wedge\theta'$. Thus $\varrho\wedge\theta=\varrho'\wedge\theta'$. Now we have

$$(\varrho\wedge\theta)_A = (\varrho'\wedge\theta')_A = \varrho'_A\wedge\theta'_A = \varrho_A\wedge\theta_A,$$

and CIP is proved. □

Summing up Propositions 3, 4, 5, we obtain:

THEOREM 6. *An algebra $\mathscr{A}$ has CEP, CIP and a complemented lattice of weak congruences if and only if it satisfies conditions* 1), 2) *and* 3) *from Proposition* 3. □

Since modularity of $C_w(\mathscr{A})$ implies CEP and CIP (see [1]), Theorem 6 yields the following

COROLLARY 7. *For an algebra $\mathscr{A}$ with a modular lattice of weak congruences,* 1), 2) *and* 3) *are necessary and sufficient conditions in order that the lattice $C_w(\mathscr{A})$ be complemented.* □

The last step in our considerations is to investigate when $C_w(\mathscr{A})$ is a Boolean lattice.

THEOREM 8. *Let $\mathscr{A}$ be an algebra. Then $C_w(\mathscr{A})$ is a Boolean lattice if and only if $\mathscr{A}$ satisfies*

1) *(see Proposition 3);*
4) *$C(\mathscr{A})$ is a Boolean lattice;*
5) *$S(\mathscr{A})$ is a Boolean lattice.*

PROOF. Let $C_w(\mathscr{A})$ be a Boolean lattice. Then 1) holds by (II) and Proposition 3. $S(\mathscr{A})$ and $C(\mathscr{A})$ are complemented (again by Proposition 3) and being distributive (as sublattices of $C_w(\mathscr{A})$) they are Boolean as well.

On the other hand, if Conditions 1), 4) and 5) are satisfied, then $C_w(\mathcal{A})$ is isomorphic with the direct product of the lattices $S(\mathcal{A})$ and $C(\mathcal{A})$. This isomorphism is established by the mapping

$$h\colon C_w(\mathcal{A}) \to S(\mathcal{A})\times C(\mathcal{A})$$

defined by $h(\varrho)=(\mathcal{B}, \varrho_A)$ where $\varrho\in C(\mathcal{B})$. Indeed, h is a bijection by (1) and it obviously preserves the order. We shall prove that h^{-1} also preserves the order. Let $(\mathcal{B}, \bar{\varrho})\leqq(\mathcal{C}, \bar{\theta})$. Then $\mathcal{B}\leqq\mathcal{C}\leqq\mathcal{A}$ and $\bar{\varrho}<\bar{\theta}$, where $\bar{\varrho}, \bar{\theta}\in C(\mathcal{A})$. By 1) take the unique $\varrho\in C(\mathcal{B})$ and $\theta\in C(\mathcal{C})$ such that $\varrho_A=\bar{\varrho}$ and $\theta_A=\bar{\theta}$. Then

$$(\varrho_C)_A = \varrho_A = \bar{\varrho} \leqq \bar{\theta} = \theta_A,$$

hence $\varrho_C\leqq\theta$. Therefore $\varrho\leqq\varrho_C\leqq\theta$, and we are done. □

EXAMPLE. Let G be the groupoid given by the Cayley table below, with constant b; this algebra has the lattice of weak congruences $C_w(G)$ shown in the figure. This lattice is Boolean, and G satisfies the properties listed in Theorem 8.

	a	b	c
a	a	b	b
b	a	a	c
c	a	c	c

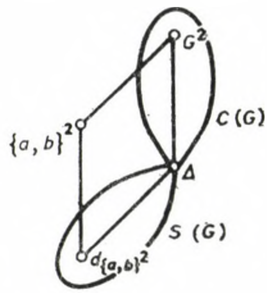

Fig. 1

The authors express their thanks to L. Márki for his valuable suggestions during the preparation of the final version of the paper.

REFERENCE

[1] VOJVODIĆ, G. and ŠEŠELJA, B., On the lattice of weak congruence relations, *Algebra Universalis* **25** (1988), 121—130.

(Received October 12, 1986)

INSTITUTE OF MATHEMATICS
UNIVERSITY OF NOVI SAD
DR ILIJE DJURIČIĆA 4
YU—21000 NOVI SAD
YUGOSLAVIA

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 295—302

ON O-SEMIGROUPS AND O*-SEMIGROUPS

SUSAN H. GENSEMER and HANNS J. WEINERT

§ 1. Introduction

In the following, $(S, \cdot)$ always denotes a semigroup, and we write $(S, \cdot)\subseteq(T, \cdot)$ iff $(S, \cdot)$ is a subsemigroup of $(T, \cdot)$. Even for sets $S\subseteq T$, a partial order (p.o.) $\leqq_T$ on T is called an *extension* of a p.o. $\leqq_S$ on S iff

(1.1) $$a \leqq_S b \Rightarrow a \leqq_T b \quad \text{holds for all} \quad a, b\in S.$$

In particular, $\leqq_T$ is called a *strict extension* of $\leqq_S$ iff

(1.2) $$a \leqq_S b \Leftrightarrow a \leqq_T b \quad \text{holds for all} \quad a, b\in S,$$

i.e. iff $\leqq_S$ is the restriction of $\leqq_T$ on S. Moreover, we speak about a *monotone partial order (m.p.o.)* $\leqq_S$ on $(S, \cdot)$ iff $(S, \cdot, \leqq_S)$ is a partially ordered (p.o.) semigroup, and correspondingly about a *monotone full order (m.f.o.)*. Clearly, if $\leqq_T$ is a m.p.o. or a m.f.o. on $(T, \cdot)$, the same holds for its restriction on each $(S, \cdot)\subseteq(T, \cdot)$.

The main subject of this paper is given by the following two concepts for semigroups, which generalize well known ones for groups (cf. [1], Chap. III):

Definition 1.1. A semigroup $(S, \cdot)$ is called on *O-semigroup* iff there exists at least one m.f.o. on $(S, \cdot)$. We call $(S, \cdot)$ an *O*-semigroup* iff each m.p.o. $\leqq_S$ on $(S, \cdot)$ has an extension $\leqq_S'$ which is a m.f.o. on $(S, \cdot)$. Since each m.p.o. on a semigroup $(S, \cdot)$ can be extended to a maximal one, the latter means that each maximal m.p.o. on $(S, \cdot)$ is a m.f.o.

As usual, we consider the equality on S as a m.p.o. on $(S, \cdot)$. Thus each O*-semigroup is also an O-semigroup, whereas the converse fails to be true even in the case of groups (cf. [1'], III. Satz 19, IV. Satz 8).

From considerations concerning fully ordered semigroups various results on O-semigroups are known, and we refer to [2] also for further references. In §2 we compare some of these results with those on O*-semigroups, which also yields some initial examples. In particular, we obtain all varieties of semigroups which contain only O*-semigroups (Prop. 2.3.).

1980 *Mathematics Subject Classification* (1985 Revision). Primary 06F05; Secondary 20M10, 06F15, 20F60, 20M14.

Key words and phrases. Partially and fully ordered semigroups, extensions of partial orders within a semigroup, in particular to full orders, O-semigroups and O*-semigroups, O-groups and O*-groups, extensions of partial orders to oversemigroups, in particular to semigroups of right quotients.

In § 4 we consider the situation where $(T, \cdot)$ is a semigroup of right quotients of a semigroup $(S, \cdot)$ with respect to a subsemigroup $(\Sigma, \cdot)$. In this case, $(S, \cdot)$ is an O-semigroup iff $(T, \cdot)$ has this property, and if $(S, \cdot)$ is an O*-semigroup, the same holds for $(T, \cdot)$ (Thm. 4.3). We could not decide whether the converse of the latter holds in this generality, but we can prove it if $(T, \cdot)$ is both, a semigroup of right and left quotients of $(S, \cdot)$ with respect to $(\Sigma, \cdot)$ (Thm. 4.4). These considerations are prepared in §3, where we deal with the extension of a m.p.o. $\leqq_S$ on $(S, \cdot)$ to a m.p.o. $\leqq_T$ on $(T, \cdot)$ for some $(T, \cdot)\supseteq(S, \cdot)$ in a more general setting.

The statements of § 4 allow us to apply results on O- or O*-groups to semigroups $(S, \cdot)$ which have a group $(T, \cdot)$ of right (and left) quotients. In this way we obtain (Thm. 5.2) that a commutative and cancellative semigroup $(S, \cdot)$ is an O-semigroup iff it is an O*-semigroup, which in turn is the case iff $(S, \cdot)$ is power cancellative (cf. §2). Whereas a semigroup of this kind is also torsion free, the converse need not be true. In this context we have to correct a wrong statement given in [7], which is done at the end of § 5. The authors are grateful to the referee for several hints and corrections.

§ 2. Some examples and general statements

LEMMA 2.1. *Let $(S, \cdot)$ be a semigroup generated by one element $a\in S$. Then the following statements are equivalent:*

a) *$(S, \cdot)$ is an O-semigroup.*

b) *$(S, \cdot)$ is an O*-semigroup.*

c) *a has infinite order or $a^m=a^{m+1}$ holds for some $m\in\mathrm{N}$.*

PROOF. It is known that a)⇔c) holds (cf. [2], § 0). In this case, there are two m.f.o. on $(S, \cdot)$, given by $a^i\leqq a^{i+k}$ for all $i, k\in\mathrm{N}$ and the dual order. We show that each such O-semigroup is also an O*-semigroup. Assume at first $a^m=a^{m+1}$. We may choose m minimal and assume $1<m$. Then there is no m.p.o. $\leqq_S$ on $(S, \cdot)$ satisfying $a^i<_S a^{i+k}$ and $a^{j+l}<_S a^j$ for some $i, j, k, l\in\mathrm{N}$. Indeed, the first relation would imply $a^{m-1}<_S a^m$ and the second one $a^m<_S a^{m-1}$. Hence each m.p.o. on $(S, \cdot)$ can be extended to one of the two m.f.o. on $(S, \cdot)$. If a has infinite order one can proceed similarly: choosing $i=j$ one obtains the contradiction $a^i<_S a^{i+k\cdot l}<_S a^i$. But in this case $(S, \cdot)$ is a commutative and cancellative O-semigroup and we can also apply Theorem 5.2 to obtain b).

It is further obvious that each subsemigroup of an O-semigroup $(T, \cdot)$ is also an O-semigroup. Consequently, each element $a\in T$ satisfies c) of Lemma 2.1. Contrasting the former statement, we state:

LEMMA 2.2. *A subsemigroup $(S, \cdot)$ of an O*-semigroup $(T, \cdot)$ need not be an O*-semigroup.*

PROOF. As an example, consider the subsemigroup $S=\{a^2, \ldots, a^6\}$ of the O*-semigroup $T=\{a^1, \ldots, a^6=a^7\}$. Then $a^4<_S a^5$ and $a^6<_S a^5$ determine obviously a m.p.o. $\leqq_S$ on $(S, \cdot)$. The latter cannot be extended to a m.f.o. $\leqq$ on $(S, \cdot)$, since $a^2<a^3$ would contradict $a^6<_S a^5$, whereas $a^3<a^2$ would contradict $a^4<_S a^5$.

In [3] a complete list is given of those varieties of semigroups which consist entirely of O-semigroups. They are defined by one of the following laws: $xy=xz$, $yx=zx$, $xy=x$, $xy=y$, $xy=uv$, and $x=y$. Using this result we obtain:

PROPOSITION 2.3. *The six varieties given above are all varieties of semigroups which consist entirely of O*-semigroups.*

PROOF. If $(S, \cdot)$ is a semigroup contained in one of the last four varieties, clearly each f.o. $\leqq_S$ is monotone on $(S, \cdot)$ (cf. [3]). This yields that each m.p.o. on $(S, \cdot)$ can be extended to a m.f.o. So it remains to prove that each semigroup $(S, \cdot)$ contained in one of the other varieties, say in that defined by $xy=xz$, is an O*-semigroup.

Let $(S, \cdot)$ be such a semigroup. Then each element $a\in S$ is either idempotent, which yields $as=a$ for all $s\in S$, or $a\notin S^2$ holds and hence $as=a^2$ for all $s\in S$. We show that each maximal m.p.o. $\leqq$ on $(S, \cdot)$ is a full order (cf. Def. 1.1).

Assume at first that a^2 and b^2 would be incomparable with respect to $\leqq$ for some $a, b\in S$. Then one checks in a straightforward manner that

$$x \leqq^* y \quad \text{iff} \quad x \leqq y \quad \text{or} \quad x \leqq a^2,\ b^2 \leqq y$$

would define a m.p.o. on $(S, \cdot)$ Since $\leqq^*$ would be a proper extension of $\leqq$, we obtain that all $a^2, b^2\in S$ are comparable with respect to $\leqq$.

Now assume that $a, b\in S$ would be incomparable with respect to $\leqq$. As just proved we may assume that $a^2\leqq b^2$ holds and obtain a contradiction like above using

$$x \leqq^* y \quad \text{iff} \quad x \leqq y \quad \text{or} \quad x \leqq a,\ b \leqq y.$$

Whereas the direct product of O-groups is again an O-group and that of O*-groups even an O*-group (cf. [1′]), the situation for semigroups differs considerably:

LEMMA 2.4. *The direct product of two O*-semigroups $(S_i, \cdot)$ need not even be an O-semigroup.*

PROOF. For an example, let $S_1=\{a, b\}$ be a left and $S_2=\{1, 2\}$ be a right zero semigroup. Both are O*-semigroups by Prop. 2.3. We write a_1, a_2, b_1, b_2 for the elements of their direct product $(T, \cdot)$ according to $x_iy_j=x_j$. By way of contradiction, let $\leqq_T$ be a m.f.o. on $(T, \cdot)$. Without loss of generality, we may assume that $a_1<_T b_2$ holds. This yields $a_1<_T a_2$ and $a_1<_T b_1$. Now $a_2<_T b_1$ implies the contradiction $a_2<_T a_1$, and $b_1<_T a_2$ the contradiction $b_1<_T a_1$.

Finally, we need two more concepts in our context. A semigroup $(S, \cdot)$ is called *power cancellative* (cf. [5], III. 6.8.) iff, for all $a, b\in S$ and some $n\in \mathbf{N}$, $a^n=b^n$ implies $a=b$. We further generalize a well known group theoretical concept and call $(S, \cdot)$ *torsion free* iff all elements $a\in S$, except the identity if there is one, have infinite order.

LEMMA 2.5. *Let $(S, \cdot)$ be cancellative and power cancellative. Then $(S, \cdot)$ is torsion free.*

PROOF. Let $a\in S$ be of finite order. Then $a^n=a^{n+k}$ holds for some $n, k\in\mathbf{N}$. This yields $a^na^k=a^n(a^k)^2$, hence $a^k=(a^k)^2$ by cancellativity. As an idempotent in a cancellative semigroup, a^k is the identity e of $(S, \cdot)$. Now $a^k = e = e^k$ implies $a = e$ by power cancellativity, i.e. $(S, \cdot)$ is torsion free.

LEMMA 2.6. *Each left cancellative O-semigroup $(S, \cdot)$ is power cancellative.*

PROOF. Assume $a^n = b^n$ for some $a, b \in S$. Then we may choose $n \in \mathbf{N}$ to be minimal and, by way of contradiction, $a <_S b$ for a m.f.o. $\leqq_S$ on $(S, \cdot)$. This yields $a^{n-1} <_S b^{n-1}$, hence $aa^{n-1} <_S ab^{n-1}$ by left cancellativity. From $ab^{n-1} \leqq_S bb^{n-1}$ we obtain the contradiction $a^n <_S b^n$.

By the last two statements, each cancellative O-semigroup is power-cancellative and torsion free, whereas Lemma 2.1 provides commutative O*-semigroups without these properties. Moreover, there are left cancellative and hence power cancellative O*-semigroups, which are not torsion free. Close at hand examples are right zero semigroups (cf. Prop. 2.3), but there are also O*-semigroups of this kind consisting of idempotents and elements of infinite order.

§ 3. Monotone partial orders on semigroups and subsemigroups

LEMMA 3.1. a) *For $(S, \cdot) \subseteq (T, \cdot)$, let $\leqq_S$ be a m.p.o. on $(S, \cdot)$. If $\leqq_S$ can be extended to a m.p.o. $\leqq_T$ on $(T, \cdot)$, then $\leqq_S$ can be extended to a m.p.o. $\leqq'_S$ on $(S, \cdot)$ such that $\leqq_T$ is a strict extension of $\leqq'_S$.*

b) *Assume additionally that $(T, \cdot)$ has an identity and let Σ be a subsemigroup of $(S, \cdot)$ for which each $\xi \in \Sigma$ has an inverse $\xi^{-1} \in T$. Then $\leqq'_S$ satisfies*

$$\text{(3.1)} \qquad \xi a \eta <'_S \xi b \eta \Rightarrow a \leqq'_S b \textit{ for all } \quad a, b \in S; \quad \xi, \eta \in \Sigma.$$

PROOF. Clearly, the restriction $\leqq'_S$ of $\leqq_T$ on S is the unique m.p.o. on $(S, \cdot)$ satisfying a), which implies b) since (3.1) holds in $(T, \cdot, \leqq_T)$.

Now let Σ be any subsemigroup of $(S, \cdot)$. If a m.p.o. $\leqq_S$ on $(S, \cdot)$ can be extended to a m.p.o. $\leqq'_S$ on $(S, \cdot)$ satisfying (3.1), then there clearly exists a unique smallest extension $\leqq'_S$ of this kind. (Note that there are those extensions $\leqq'_S$ if $\leqq_S$ can be extended to a m.f.o. on $(S, \cdot)$, hence if $(S, \cdot)$ is an O*-semigroup.) In this context and in order to use Lemma 3.1 to construct an extension $\leqq_T$ of $\leqq_S$ via $\leqq'_S$ in the next section we state:

LEMMA 3.2. *Let $\leqq_S$ be a m.p.o. on $(S, \cdot)$ and $(\Sigma, \cdot) \subseteq (S, \cdot)$ such that each $\xi \in \Sigma$ is cancellable in $(S, \cdot)$, Then*

$$\text{(3.2)} \qquad a \leqq'_S b \Leftrightarrow \xi a \eta \leqq_S \xi b \eta \quad \textit{for some} \quad \xi, \eta \in \Sigma$$

defines a m.p.o. $\leqq'_S$ on $(S, \cdot)$ which extends $\leqq_S$ and satisfies (3.1) *iff the following condition holds for all $a, b \in S$ and all $\xi, \eta \in \Sigma$:*

$$\text{(3.3)} \qquad \xi a \eta \leqq_S \xi b \eta \Rightarrow \forall x, y \in S^1 \, \exists \varkappa, \lambda \in \Sigma \quad \textit{such that} \quad \varkappa x a y \lambda \leqq_S \varkappa x b y \lambda.$$

If this is the case, clearly $\leqq'_S$ is the smallest extension of $\leqq_S$ which satisfies (3.1) *for Σ.*

PROOF. If (3.2) defines such an extension $\leqq'_S$ of $\leqq_S$,

$$\xi a\eta \leqq_S \xi b\eta \Rightarrow a \leqq'_S b \Rightarrow xay \leqq'_S xby \Rightarrow \varkappa xay\lambda \leqq_S \varkappa xby\lambda$$

shows that (3.3) holds. Conversely, assume (3.3) and define a relation $\leqq'_S$ by (3.2). Clearly, $\leqq'_S$ extends $\leqq_S$ and is hence reflexive, too. To check transitivity, assume $a\leqq'_S b$ and $b\leqq'_S c$, i.e. $\xi a\eta \leqq_S \xi b\eta$ and $\sigma b\tau \leqq_S \sigma c\tau$ for some $\xi, \eta, \sigma, \tau\in\Sigma$. Applying (3.3) for $x=\sigma$ and $y=\tau$, we get

$$\varkappa\sigma a\tau\lambda \leqq_S \varkappa\sigma b\tau\lambda \leqq_S \varkappa\sigma c\tau\lambda$$

and hence $a \leqq'_S c$ because of $\varkappa\sigma, \tau\lambda\in\Sigma$. The same calculation with $c=a$ yields $\varkappa\sigma a\tau\lambda = \varkappa\sigma b\tau\lambda$, hence $a=b$ since $\varkappa\sigma$ and $\tau\lambda$ are cancellable in $(S, \cdot)$. So $\leqq'_S$ is also antisymmetric, thus a p.o. on S. The monotony for $\leqq'_S$ follows immediately from (3.3). Finally, $\xi a\eta <'_S \xi b\eta$ implies $\sigma\xi a\eta\tau \leqq_S \sigma\xi b\eta\tau$ for some $\sigma, \tau\in\Sigma$ and hence $a\leqq'_S b$, i.e. $\leqq'_S$ satisfies (3.1).

We close this section with the following

EXAMPLE 3.3. Let $(T, \cdot)$ be the group of all matrices

$$A = \begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} \quad \text{for} \quad a, b, c\in\mathbf{Z} \quad (\text{or } \mathbf{Q}, \text{ or } \mathbf{R})$$

with the usual multiplication. We use the abbreviation $A=(a, b, c)$ and state

$$A\cdot X = (a, b, c)(x, y, z) = (a+x, b+y, c+ay+z),$$
$$A^{-1} = (a, b, c)^{-1} = (-a, -b, ab-c).$$

Obviously, $Z_1=\{(0, 0, c)|c\in\mathbf{Z}\}$ is the centre of $(T, \cdot)$ and T/Z_1 is commutative, hence $(T, \cdot)$ a nilpotent group. Since $(T, \cdot)$ is also torsion free, it is an O*-group by [1′], III. Satz 17. For the subsemigroup given by

$$S = \{(a, b, c)|a \geqq 0, \ b \geqq 0, \ c \geqq 0\}$$

with respect to the usual m.f.o. $\leqq$ on $\mathbf{Z}$, one easily checks that a m.p.o. on $(S, \cdot)$ is defined by

(3.4) $\qquad (a, b, c) \leqq_S (a', b', c') \Leftrightarrow a < a' \quad \text{or} \quad a = a', \ b \leqq b', \ c \leqq c'.$

Note that (3.4) does not define a m.p.o. on $(T, \cdot)$, since $(0, 2, 0) <_S (0, 3, 0)$ multiplied by $(-1, 0, 0)$ from the left yields $(-1, 2, -2)$ on the left and $(-1, 3, -3)$ on the right side. Moreover, choosing $\Sigma=S$, the m.p.o. $\leqq_S$ on $(S, \cdot)$ does not satisfy (3.1). A simple counterexample is

$$(1, 0, 0)(0, 1, 2)(0, 0, 0) = (1, 1, 3) <_S (1, 2, 3) = (1, 0, 0)(0, 2, 1)(0, 0, 0),$$

whereas $(0, 1, 2)$ and $(0, 2, 1)$ are incomparable with respect to $\leqq_S$. It is straightforward to check that (3.2) applied to $\leqq_S$ defines a relation $\leqq'_S$ on S given by

(3.5) $(a, b, c) \leqq'_S (a', b', c') \Leftrightarrow a < a' \quad \text{or} \quad a = a', b < b' \quad \text{or} \quad a = a', b = b, c \leqq c'.$

Clearly, $\leqq'_S$ is even a m.f.o. on $(S, \cdot)$ which extends $\leqq_S$ and satisfies (3.1). The latter implies that (3.3) holds for $\leqq_S$. Contrasting the situation with (3.4), (3.5) defines also a m.p.o. $\leqq_T$ on $(T, \cdot)$ which is, of course, a m.f.o. and a strict extension of $\leqq'_S$. We shall pick up this example in the next section; in particular we obtain there directly that $\leqq_S$ satisfies (3.3).

§ 4. O- and O*-semigroups and extensions by quotients

Assume $(\Sigma, \cdot)\subseteqq(S, \cdot)\subseteqq(T, \cdot)$. Then T is called a *semigroup of right quotients (s.o.r.q.) of S with respect to* Σ iff T has an identity, each $\alpha\in\Sigma$ has an inverse $\alpha^{-1}\in T$ and the subset $\{a\alpha^{-1}|a\in S, \alpha\in\Sigma\}$ of T coincides with T. It is well known (cf. e.g. [6]) that such a semigroup T exists iff each $\alpha\in\Sigma$ is cancellable in S and the following condition, denoted by $q_r(S, \Sigma)$, is satisfied

$$(q_r(S, \Sigma)) \qquad\qquad a\Sigma\cap\alpha S \neq \emptyset \quad \text{for all} \quad a\in S,\ \alpha\in\Sigma.$$

In this case, T is uniquely determined by S and Σ (up to isomorphisms relative with respect to S), hence we may write $T=Q_r(S, \Sigma)$. We need a statement proved in [6] (Thm. 2) and reformulate it as follows:

THEOREM 4.1. *Let $T=Q_r(S, \Sigma)$ be a s.o.r.q. and $\leqq'_S$ a m.p.o. on $(S, \cdot)$. Then there exists a m.p.o. $\leqq_T$ on $(T, \cdot)$ which is a strict extension of $\leqq'_S$ iff $\leqq'_S$ satisfies* (3.1). *If this is the case, $\leqq_T$ is uniquely determined by $\leqq'_S$ according to*

$$(4.1) \qquad a\alpha^{-1} \leqq_T b\beta^{-1} \Leftrightarrow \alpha\lambda = \beta l \text{ and } a\lambda \leqq'_S bl \text{ for some } l\in S, \lambda\in\Sigma,$$

and $\leqq_T$ is a full order iff $\leqq'_S$ is one.

COROLLARY 4.2. *For $T=Q_r(S, \Sigma)$ as above, a m.f.o. $\leqq'_S$ on $(S, \cdot)$ can always be extended to a m.p.o. $\leqq_T$ on $(T, \cdot)$. Moreover, $\leqq_T$ is also a full order, uniquely determined by $\leqq'_S$, and a strict extension of $\leqq'_S$.*

PROOF. Since (3.1) holds for each m.f.o. $\leqq'_S$ on $(S, \cdot)$, there exists the m.f.o. $\leqq_T$ of $(T, \cdot)$ given by (4.1), which is a strict extension of $\leqq'_S$. We show that each m.p.o. on $(T, \cdot)$ which extends $\leqq'_S$, say $\leqq^*_T$, coincides with $\leqq_T$. Assume at first $a\alpha^{-1} \leqq^*_T b\beta^{-1}$ for some $a\alpha^{-1}, b\beta^{-1}\in T$. By $q_r(S, \Sigma)$ there are $l\in S$, $\lambda\in\Sigma$ satisfying $\alpha\lambda=\beta l$. This yields $a\lambda \leqq^*_T bl$ and hence $a\lambda \leqq'_S bl$, since $a\lambda >'_S bl$ would imply $a\lambda >^*_T bl$. Conversely, assume $\alpha\lambda=\beta l$ and $a\lambda \leqq'_S bl$. This yields $l\lambda^{-1}=\beta^{-1}\alpha$ and $a\lambda\leqq^*_T bl$, hence

$$a\alpha^{-1} = a\lambda\lambda^{-1}\alpha^{-1} \leqq^*_T bl\lambda^{-1}\alpha^{-1} = b\beta^{-1}\alpha\alpha^{-1} = b\beta^{-1}.$$

So $\leqq^*_T$ coincides with $\leqq_T$ according to (4.1).

THEOREM 4.3. *Let $T=Q_r(S, \Sigma)$ be a s.o.r.q. Then*
a) $(S, \cdot)$ *is an O-semigroup iff $(T, \cdot)$ is an O-semigroup and*
b) *if $(S, \cdot)$ is an O*-semigroup, the same holds for $(T, \cdot)$.*

PROOF. Since a subsemigroup of an O-semigroup is an O-semigroup, a) follows from Cor. 4.2. For b) let $\leqq_T$ be a m.p.o. on $(T, \cdot)$. Then its restriction $\leqq'_S$ on $(S, \cdot)$ can be extended to a m.f.o. $\leqq''_S$ on $(S, \cdot)$ which can be extended to a m.f.o. $\leqq''_T$

on $(T, \cdot)$ by Cor. 4.2. To see that $\leqq''_T$ extends $\leqq_T$ assume $a\alpha^{-1} \leqq_T b\beta^{-1}$. By $q_r(S, \Sigma)$, this implies $\alpha\lambda=\beta l$ and $a\lambda \leqq_T bl$ for some $l \in S$ and $\lambda \in \Sigma$, hence $a\lambda \leqq'_S bl$ and $a\lambda \leqq''_S bl$. So we obtain $a\alpha^{-1} \leqq''_T b\beta^{-1}$ according to (4.1).

Concerning the converse of Thm. 4.3 b), assume that $(T, \cdot)$ is an O*-semigroup. Then $(S, \cdot)$ is also an O*-semigroup iff each m.p.o. $\leqq_S$ on $(S, \cdot)$ can be extended to a m.p.o. $\leqq'_S$ on $(S, \cdot)$ which satisfies (3.1), where the non-trivial implication follows from Thm. 4.1. Unfortunately, we need Lemma 3.2 and thus (3.3) to prove the existence of $\leqq'_S$. Considering (3.3), for all $\eta \in \Sigma$, $y \in S$ there are $l \in S$ and $\lambda \in \Sigma$ such that $\eta l = y\lambda$ holds, which is just the condition $q_r(S, \Sigma)$. The left-right dual one, $q_l(S, \Sigma)$, means that for all $\xi \in \Sigma$, $x \in S$ there are $k \in S$ and $\varkappa \in \Sigma$ satisfying $k\xi = \varkappa x$. Thus both together imply (3.3) in multiplying $\xi a\eta \leqq_S \xi b\eta$ by k from the left and by l from the right. So, using also the dual concept of a *semigroup of left quotients (s.o.l.q.) $Q_l(S, \Sigma)$ of S with respect to Σ*, we have shown:

THEOREM 4.4. *Let $T=Q_r(S, \Sigma)=Q_l(S, \Sigma)$ be a s.o.r.q. and a s.o.l.q. Then $(S, \cdot)$ is an O*-semigroups iff $(T, \cdot)$ is an O*-semigroup.*

REMARK 4.5. For $(\Sigma, \cdot) \subseteqq (S, \cdot)$, assume that the elements of Σ are cancellable in $(S, \cdot)$ and that both conditions, $q_r(S, \Sigma)$ and $q_l(S, \Sigma)$, are satisfied. Then, clearly, the (essentially unique) s.o.r.q. $T=Q_r(S, \Sigma)$ is also a s.o.l.q., and vice versa. This is just the situation $T=Q_r(S, \Sigma)=Q_l(S, \Sigma)$ assumed in Thm. 4.4, and we can speak about the *semigroup of quotients (s.o.q.) $T=Q(S, \Sigma)$ of S with respect to Σ* in this case. In particular, $q_r(S, \Sigma)$ and $q_l(S, \Sigma)$ are trivially satisfied if Σ is contained in the centre of $(S, \cdot)$.

Finally, we consider the semigroup $(S, \cdot)$ and the group $(T, \cdot)$ of Expl. 3.3, choosing $\Sigma = S$. One easily checks that for each element $C \in T$ there are elements $A, X, B, Y \in S$ satisfying $C = AX^{-1} = Y^{-1}B$. So we have $T=Q_r(S, S)=Q_l(S, S)$ or $T=Q(S, S)$. Since $(T, \cdot)$ was shown to be an O*-semigroup, we obtain that $(S, \cdot)$ is an O*-semigroup by Thm. 4.4. For the same reason, each semigroup $(H, \cdot)$ satisfying $(S, \cdot) \subseteqq (H, \cdot) \subseteqq (T, \cdot)$ is also an O*-semigroup.

§ 5. Commutative and cancellative O- and O*-semigroups

It is due to several authors (cf. [1], III. Cor. 5 and Cor. 13) that a commutative group $(G, \cdot)$ is an O-Group as well as an O*-group iff $(G, \cdot)$ is torsion free. On the other hand, each commutative and cancellative semigroup $(S, \cdot)$ is contained in a commutative group, where the smallest one is the group $(G, \cdot)=Q(S, S)$ of quotients of S (cf. Remark 4.5). So, by Thms. 4.3 and 4.4, $(S, \cdot)$ is an O-semigroup or an O*-semigroup iff $(G, \cdot)=Q(S, S)$ has the same property. To combine these results we state:

LEMMA 5.1. *Let $(S, \cdot)$ be commutative and cancellative and $G=Q(S, S)$ the group of quotients of S. Then $(G, \cdot)$ is torsion free iff $(S, \cdot)$ is power cancellative.*

PROOF. For all $a, b \in S$ and integers $m=n+k>n>0$ one has

$$(ab^{-1})^m = (ab^{-1})^n \Leftrightarrow a^m b^n = a^n b^m \Leftrightarrow a^k = b^k.$$

Now the left-hand equation implies $a=b$ iff $(G, \cdot)$ is torsion free, whereas $a^k=b^k\Rightarrow$ $\Rightarrow a=b$ holds iff $(S, \cdot)$ is power cancellative. This yields our statement.

THEOREM 5.2. *For each commutative and cancellative semigroup* $(S, \cdot)$ *the following statements are equivalent:*

a) $(S, \cdot)$ *is an O-semigroup.*

b) $(S, \cdot)$ *is an* O^**-semigroup.*

c) $(S, \cdot)$ *is power cancellative.*

PROOF. We already know b)$\Rightarrow$a)$\Rightarrow$c), the latter by Lemma 2.6. Now c) implies by Lemma 5.1 that $G=Q(S, S)$ is torsion free and hence an O*-group as noted above. This yields by Thm. 4.4 that $(S, \cdot)$ is an O*-semigroup, i.e. b).

As announced in the introduction, we have to correct an error in this context. In the corollary of Thm. 1 c in [7] it was claimed that, using the notion of this paper, a commutative and cancellative semigroup $(S, \cdot)$ is an O*-semigroup iff $(S, \cdot)$ is torsion free. According to Thm. 5.2, the latter has to be replaced by "power cancellative", since there are commutative and cancellative semigroups which are torsion free, but not power cancellative (cf. Lemma 2.5). A simple example is the semigroup $(S, \cdot)=(2\mathbf{Z}\setminus\{0\}, \cdot)$ of non-zero even integers (e.g. $(-2)^2=2^2$). Note that its group of quotients $G=Q(S, S)$, clearly the non-zero rational numbers, is not torsion free since $-1\in G$ has order 2.

REFERENCES

[1] FUCHS, L., *Partially ordered algebraic systems*, Pergamon Press, Oxford, 1963. *MR* **30** #2090.

[1'] FUCHS, L., *Teilweise geordnete algebraische Strukturen*, Studia Mathematica-Mathematische Lehrbücher, Bd. XIX, Vandenhoeck & Ruprecht, Göttingen, 1966. *MR* **34** #4386.

[2] GABOVIČ, E. JA., Totally ordered semigroups and their applications, *Uspehi Mat. Nauk* **31** (1976), no. 1 (187), 137—201 (in Russian). *Zbl* **335**.06015; *MR* **58** #21886. Translated in *Russian Math. Surveys* **31** (1976), no. 1, 147—216.

[3] KUSPANOV, U. K., Varieties of semigroups whose semigroups are all orderable, Proc. Twelfth All-Union Algebra Coll., Vol. 2, 213, Sverdlovsk, 1973 (in Russian).

[4] MALCEV, A. I., On the completion of group order, *Trudy Mat. Inst. Steklov* **38** (1951), 173—175. *MR* **14**—13.

[5] PETRICH, M., *Lectures in semigroups*, John Wiley and Sons, London, 1977. *MR* **57** #6150.

[6] WEINERT, H. J., On the extensions of partial orders on semigroups of right quotients, *Trans. Amer. Math. Soc.* **142** (1969), 345—353. *MR* **39** #6801.

[7] WEINERT, H. J., Extension of partial orders on a semigroup to semigroups of quotients, *Acta Math. Acad. Sci. Hungar.* **21** (1970), 449—451. *MR* **42** #4462.

(Received November 25, 1986)

DEPARTMENT OF ECONOMICS
SYRACUSE UNIVERSITY
SYRACUSE, NY 13210
U.S.A.

INSTITUT FÜR MATHEMATIK
TECHNISCHE UNIVERSITÄT CLAUSTHAL
ERZSTRASSE 1
D—3392 CLAUSTHAL-ZELLERFELD
FEDERAL REPUBLIC OF GERMANY

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 303—304

A COUNTEREXAMPLE TO AN ISOPERIMETRIC PROBLEM OF L. FEJES TÓTH

G. KERTÉSZ

We recall a theorem of G. Fejes Tóth [1]:

Decompose a convex polygon of at most six sides into N convex polygons of areas $a_1, \ldots, a_N$. If $\min a_i/\max a_i$ is greater than a constant $c=0.316\ldots$ then the sum P of the perimeters of the partial polygons satisfies the inequality

$$P > \sqrt{8\sqrt{3}}\sum_{i=1}^{N}\sqrt{a_i}. \tag{1}$$

An alternative proof was given by L. Fejes Tóth [2] who made the "tentative conjecture" that the theorem continues to hold without the stipulation that $\min a_i/\max a_i > c$. In the present paper we disprove this conjecture by constructing a periodical tiling of the plane with convex tiles of average perimeter smaller than the average perimeter of regular hexagons with the same areas as the tiles.

We "truncate" the vertices of a regular hexagonal tiling of unit side-length by equilateral triangles of side-length x. Decomposing each triangle into an equiangular hexagon and three equilateral triangles of side-length xy we obtain a tiling $T(x, y)$ with triangular, hexagonal and dodecagonal tiles. One dodecagon, two hexagons abutting on opposite sides of the dodecagon and the six triangles abutting on these two hexagons form a period of $T(x, y)$.

Let $a_3=\sqrt{3}x^2y^2/4$, $a_6=\sqrt{3}x^2(1-3y^2)/4$ and $a_{12}=\sqrt{3}(3-x^2)/2$ be the area of the trigonal, hexagonal and dodecagonal tiles. Let $p=6+(12-4\sqrt{3}+12y)x$ be the total perimeter of tiles in a period and let $h=\sqrt{8\sqrt{3}}\,(12\sqrt{a_3}+2\sqrt{a_6}+\sqrt{a_{12}})$ be the sum of perimeters of hexagons with the same areas as the tiles in a period.

A counterexample to the above conjecture will be given by finding values $0<x, y<1/2$ such that $p/h<1$. Among such values we are looking for those

(i) for which p/h attains its minimum,

(ii) for which a_3/a_{12} attains its maximum.

1980 *Mathematics Subject Classification* (1985 Revision). Primary 52A45; Secondary 52A40.
Key words and phrases. Tiling, isoperimetric problem.

We give approximate values satisfying these conditions:

(i) $x=0{,}034$, $y=0{,}1746$, $a_3/a_{12}=5.87\ldots\cdot 10^{-6}$, $p/h=0.99981\ldots$.

(ii) $x=0.065$, $y=0.195$, $a_3/a_{12}=2.68\ldots\cdot 10^{-5}$, $p/h=0.999998\ldots$.

The question if (1) holds for a greater constant $c=0.41\ldots$ is still open. Other open problems can be found in [2].

REFERENCES

[1] Fejes Tóth, G., An isoperimetric problem for tessellations, *Studia Sci. Math. Hungar.* **10** (1975), 171—173. *MR* **56** # 3748.
[2] Fejes Tóth, L., Isoperimetric problems for tilings, *Mathematika* **32** (1985), 10—15.

(Received December 5, 1986)

MTA MATEMATIKAI KUTATÓ INTÉZETE
POSTAFIÓK 127
H—1364 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 305—306

PSHENICHNYI'S NECESSARY CONDITION FOR NONSMOOTH PROGRAMMING

CHARLES SWARTZ

Abstract

A necessary condition of Pshenichnyi for minimums in convex programming problems is generalized to non-convex minimization problems.

If f is a differentiable function on a subset E of $\mathbf{R}^m$ and f has a local minimum at the point $x\in E$, then a classical necessary condition is that $\nabla f(x)$ must belong to the dual cone of the tangent cone of E at x. Pshenichnyi has given a generalization of this result to convex functions which are not necessarily differentiable by using the subgradient ([4] II.2.1). Hiriart—Urruty has given a generalization of Pshenichnyi's result to the case of non-convex functions by using the generalized gradient of Clarke ([3] Theorem 8; [2] Theorem 3). In this brief note we make the observation that Hiriart—Urruty's generalization of the Pshenichnyi necessary condition actually follows quite directly from Pshenichnyi's condition; the proof based on Pshenyichnyi's condition is quite simple compared to the proofs given in [2], Theorem 3 and [3], Theorem 8.

To fix the notation, let X be a real Banach space with dual space X'. Let $E\subseteq X$ and $f\colon E\to\mathbf{R}$ be locally Lipschitz. Then the generalized directional derivative of f at $x\in E$ defined by Clarke is $f^0(x;d)=\varlimsup\limits_{\substack{y\to x\\ t\to 0+}}[f(y+td)-f(y)]/t$ ([1] 2.1). The function $f^0(x;\cdot)$ is the support function of a non-empty weak* compact, convex subset of X' denoted by $\partial f(x)$ and called the generalized gradient of f at x ([1], 2.1). As is indicated by the notation, the generalized gradient generalizes the subgradient of convex programming ([1] 2.2.7).

If $x\in E$, then the cone of adherent displacements for E at x is defined to be $T(E,x)=\{h\in X\colon$ there exist sequences $\{x_k\}\subseteq X$, $\{t_k\}\subseteq\mathbf{R}_+$ such that $x_k\to x$ and $h=\lim t_k(x_k-x)\}$.

We next state a necessary condition for a minimum in terms of the generalized directional derivative. This result was established by Hiriart—Urruty in [3], Theorem 6 and [2], Theorem 2.

THEOREM 1. *If f has a local minimum at $x\in E$, then $f^0(x;h)\geqq 0$ for every $h\in T(E,x)$.*

Theorem 1 is a generalization of the classical nonlinear programming result, $\nabla f(x)\cdot h\geqq 0$, for differentiable functions f defined on subsets of $\mathbf{R}^n$.

1980 *Mathematics Subject Classification.* Primary 90C30.
Key words and phrases. Generalized directional derivative, generalized gradient, cone, dual cone.

Pshenichnyi's necessary condition for a minimum in convex programs is given in terms of the dual cone of the cone of feasible directions ([4] II.2.1). We now give the generalization ([2] Theorem 3). Since the cone of adherent displacement is not in general convex, the generalization is given in terms of convex subcones of this cone. If $C \subseteq X$, then the dual cone of C is the weak* closed, convex cone given by $C^* = \{x' \in X' : \langle x', x \rangle \geqq 0$ for $x \in C\}$.

THEOREM 2. *Suppose f has a local minimum at $x \in E$ and let M be a convex subcone of $T(E, x)$. Then $\partial f(x) \cap M^* \neq \emptyset$.*

PROOF. From Theorem 1 it follows that $h=0$ solves the convex program: $\min \{f^\circ(x; h): h \in M\}$. By Pshenichnyi's necessary condition for convex programs, it follows that the subgradient $\partial[f^\circ(x; \cdot)](0)$ and the dual of the cone of feasible directions to M at 0, $F(M, 0)$, have a non-empty intersection. Since $f^0(x; \cdot)$ is the support function of $\partial f(x)$, it follows that $\partial[f(x; \cdot)](0) = \partial f(x)$ ([4] I.1.6). Since M is a convex cone, $F(M, 0) = \{e: te \in M$ for some $t > 0\} = M$ ([4] II). Thus, $\partial f(x) \cap M^* \neq \emptyset$.

In particular, Clarke's cone of tangents ([1] 2.4) is a convex cone which is always contained in the cone of adherent displacements so that Theorem 2 is applicable to this cone. Hiriart—Urruty has given an example which shows that the conclusion of Theorem 2 does not hold if M is replaced by the cone of adherent displacements ([2] p. 80).

REFERENCES

[1] CLARKE, F. H., *Optimization and nonsmooth analysis*, Canadian Mathematical Society Series of Monographs and Advanced Texts, Wiley, New York, 1983. *MR* **85m**: 49002.
[2] HIRIART-URRUTY, J. B., On optimality conditions in non-differentiable programming, *Math. Programming* **14** (1978), 73—86. *MR* **57** # 5121.
[3] HIRIART-URRUTY, J. B., Tangent cones, generalized gradients and mathematical programming in Banach spaces, *Math. Oper. Res.* **4** (1979), 79—97. *MR* **80i**: 58011.
[4] PSHENICHNYI, B. N., *Necessary conditions for an extremum*, Dekker, New York, 1971.

(Received December 19, 1986)

DEPARTMENT OF MATHEMATICAL SCIENCES
NEW MEXICO STATE UNIVERSITY
LAS CRUCES, NM 88003
U.S.A.

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 307—314

ON QUASILINEAR ELLIPTIC SYSTEMS IN R^n

K. SINGH

0. Introduction

In this paper the following system of quasilinear differential equations of order $2m$ will be considered:

$$A_j u_j + g_j(x, u(x), \ldots, D^\beta u(x), \ldots) = f_j(x), \tag{0.1}$$

where $j=1, \ldots, M$; $u(x)=(u_1, \ldots, u_M)\in \prod_{j=1}^{M} H^{2m}(R^n)$ and $|\beta|\leqq 2m-1$.

Let us suppose that

$$A_j u_j \equiv (P_j+Q_j) u_j,$$

where $P_j=P_j(D)$ is linear elliptic operator of order $2m$ with constant coefficients and $Q_j=Q_j(x, D)$ is a differential operator of order $\leqq 2m$ with smooth coefficients vanishing for $|x|>a$, $D=\left(-i\frac{\partial}{\partial x_1}, \ldots, -i\frac{\partial}{\partial x_n}\right)$. Also it will be assumed that $P_j(\xi)\neq 0$ for $\xi\in R^n\setminus\{0\}$ and g_j satisfy conditions $(a_1)-(a_3)$ of Theorem 1.1. In [3] it has been shown that when $P_j(\xi)\neq 0$ for every $\xi\in R^n$, then for every $f_j\in L_a^2(R^n)$ (i.e. $f_j\in L^2(R^n), f_j(x)=0$ if $|x|>a$) there exists a unique solution u_j of the equation

$$A_j u_j = f_j \tag{0.2}$$

which tends to zero as $|x|\to\infty$ and estimate

$$\|u_j\|_{H^{2m}(R^n)} \leqq c\|f_j\|_{L_a^2(R^n)} \tag{0.3}$$

holds. In [3], it has been formulated conditions on differential operators B^{lj} of order $<2m$ such that for the boundary value problem in $B_\varrho=\{x\in R^n: |x|<\varrho\}$, $\varrho>0$

$$(P_j+Q_j) u_j^\varrho = f_j \quad \text{in} \quad B_\varrho \tag{0.4}$$

$$B^{lj}(\omega, D)u_j^\varrho = 0 \quad \text{on} \quad S_\varrho, \quad (l = 1, \ldots, m)$$

where

$$S_\varrho = \{x\in R: |x| = \varrho\}, \omega = \frac{x}{|x|}$$

there exists a unique solution u_j^ϱ in $H^{2m}(B_\varrho)$ and

$$\|u_j-u_j^\varrho\|_{H^{2\varkappa}(B_\varrho)} \leqq c\,\|f_j\|_{L_a^2(R^n)}e^{-c'\varrho}. \tag{0.5}$$

1980 *Mathematics Subject Classification* (1985 Revision). Primary 35J60.

Key words and phrases. Partial differential equations, elliptic systems, unbounded domains.

Further, for $P_j(\xi)\neq 0$, $\xi\in R\setminus\{0\}$ and $P_j(0)=0$, it has been shown in [4] that under certain additional conditions following estimations are valid for any compact $K\subset R^n$:

$$\|u_j\|_{H^{2m}(K)} \leqq c(K)\,\|f_j\|_{L^2_a(R^n)}, \tag{0.6}$$

$$\|u_j-u_j^{\varrho}\|_{H^{2m}(K)} \leqq g(\varrho)\,\|f_j\|_{L^2_a(R^n)} \tag{0.7}$$

where $\lim\limits_{\varrho\to\infty} g(\varrho)=0$. If some other conditions are imposed then also

$$\|u_j-u_j^{\varrho}\|_{H^{2m}(B_\varrho)} \leqq g(\varrho)\,\|f_j\|_{L^2_a(R^n)} \tag{0.8}$$

holds, where $\lim\limits_{\varrho\to\infty} g(\varrho)=0$.

Problem (0.1) has been considered in [5] for a single equation. In this paper we shall prove the existence of solutions of system (0.1). A result on approximation of solutions will be given in the last section.

1. Existence theorems

THEOREM 1.1. *Let us suppose that for the equation* (0.2) *there exists a unique solution vanishing at infinity for any* $f_j\in L^2_a(R^n)$ *and the estimation* (0.6) *holds. Further assume that* $g_j\colon R^{n+NM}\to R$ *is a continuous function (N is the number of multiindices with* $|\beta|\leqq 2m-1$*) which satisfies conditions:*

$$g_j(x,\xi_0,\dots,\xi_\beta,\dots)=0 \quad \text{if} \quad |x|>a; \tag{a$_1$}$$

$$\lim_{|(\xi_0,\dots,\xi_\beta,\dots)|\to\infty} \frac{g_j(x,\xi_0,\dots,\xi_\beta,\dots)}{|(\xi_0,\dots,\xi_\beta,\dots)|}=0 \tag{a$_2$}$$

uniformly in x;

(a$_3$) *The first partial derivatives of* g_j *are continuous and bounded.*

Then for all f_j*'s belonging to* $L^2_a(R^n)$ *equation* (0.1) *must have at least one solution u which belongs to* $\prod\limits_{j=1}^{M} H^{2m}_{loc}(R^n)$ *and vanishes for* $|x|\to\infty$.

PROOF. Assume that $u_j=A_j^{-1}f_j$ is the unique solution of the equation $A_ju_j=f_j$ vanishing at infinity and let $A^{-1}=(A^{-1},\dots,A_M^{-1})$. The solution $u=(u_1,\dots,u_M)$ will be a solution of the system of equation (0.1) vanishing at infinity if and only if $v_j=A_ju_j\in\prod\limits_{j=1}^{M} L^2_a(R^n)$ is a solution of

$$v_j+G_j(v)=f_j, \qquad j=1,\dots,M \tag{1.1}$$

where $v:=(v_1,\dots,v_M)$,

$$G_j:=g_j(x,A^{-1}v,\dots,D^{\beta}A^{-1}v,\dots).$$

Let $G:=(G_1,\dots,G_M)$, $f:=(f_1,\dots,f_M)$. From estimation (0.6) it follows that the operator $A^{-1}\colon \prod\limits_{j=1}^{M} L^2_a(R^n)\to\prod\limits_{j=1}^{M} H^{2m}_a(B_a)$ is a bounded linear operator.

The assumption (a$_2$) and boundedness of A^{-1} gives that $|g_j(x, u, \ldots, D^\beta u, \ldots)| \leqq \leqq c|(u, \ldots, D^\beta u, \ldots)|$, and

$$|G_j(v)|_{L_a^2(R^n)} \leqq c\,\|A^{-1}v\|_{\prod_{j=1}^M H^{2m}(B_a)} \leqq c\,\|v\|_{\prod_{j=1}^M L_a^2(R^n)}. \tag{1.2}$$

Using the continuity and boundedness of first derivatives of g_j's and applying mean value theorem we obtain

$$|G_j(v)-G_j(v^*)| \leqq c[|A^{-1}(v-v^*)|+\ldots+|D^\beta A^{-1}(v-v^*)|+\ldots]$$

where c is a constant. Consequently,

$$\{|G_j(v)-G_j(v^*)|^2\}^{1/2} \leqq c\Big[\Big\{\int_{B_a} |A^{-1}(v-v^*)|^2\Big\}^{1/2}+\ldots+$$

$$+\ldots+\Big\{\int_{B_a} |D^\beta A^{-1}(v-v^*)|^2\Big\}^{1/2}+\ldots\Big] \leqq c\,\|v-v^*\|_{\prod_{j=1}^M L_a^2(R^n)}.$$

Thus $G\colon \prod_{j=1}^M L_a^2(R^n) \to \prod_{j=1}^M L^2(a)$ is a continuous operator. For any v belonging to $L_a^2(R^n)$, $D_k G_j(v) \in L_a^2(R^n)$. This in turn implies that $G_j(v)$ belongs to $H_a^1(R^n)$, since $G_j(v)=0$ if $|x|>a$. Further, G maps bounded sets of $\prod_{j=1}^M L_a^2(R^n)$ into bounded sets of $\prod_{j=1}^M H_a^1(R^n)$. This means that G is a compact operator over $\prod_{j=1}^M L_a^2(R^n)$.

We shall prove that

$$\lim_{\|v\|_{\prod_{j=1}^M L_a^2(R^n)} \to \infty} \frac{\|G_j(v)\|_{L_a^2(R^n)}}{\|v\|_{\prod_{j=1}^M L_a^2(R^n)}} = 0. \tag{1.3}$$

This will imply that

$$\lim_{\|v\| \to \infty} \frac{\|G(v)\|_{\prod_{j=1}^M L_a^2(R^n)}}{\|v\|_{\prod_{j=1}^M L_a^2(R^n)}} = 0.$$

Let

$$(A^{-1}v_1, \ldots, A_M^{-1}v_M) = (u_1, \ldots, u_M).$$

Thus

$$\frac{\|G_j(v)\|_{L_a^2(R^n)}}{\|v\|_{\prod_{j=1}^M L_a^2(R^n)}} = \frac{\|g_j(x, u, \ldots, D_\beta\, u, \ldots)\|_{L_a^2(R^n)}}{\|u\|_{\prod_{j=1}^M H^{2m}(B_a)}} \cdot \frac{\|u\|_{\prod_{j=1}^M H^{2m}(B_a)}}{\|v\|_{\prod_{j=1}^M L_a^2(R^n)}}.$$

Boundedness of A^{-1} implies that the second term on right-hand side is bounded and if $\|v\|_{\prod_{j=1}^M L_a^2(R^n)} \to \infty$ $\|u\|_{\prod_{j=1}^M H^{2m}(B_a)}$ will also tend to infinity since $v=A(u)$ and A is bounded linear operator.

We prove now the estimate

$$\lim_{\|u\|_{\prod_{j=1}^{M} H^{2m}(B_a)} \to \infty} \frac{\|g_j(x, u, \ldots, D^\beta u, \ldots)\|_{L_a^2(R^n)}}{\|u\|_{\prod_{j=1}^{M} H^{2m}(B_a)}} = 0. \tag{1.4}$$

Let $\delta > 0$ be an arbitrary number then

$$\int_{B_a} |g_j(x, u(x), \ldots, D^\beta u(x), \ldots)|^2 \, dx =$$

$$= \int_{|(u(x), \ldots, D^\beta u(x), \ldots)| > \delta} |g_j(x, u(x), \ldots, D^\beta u(x), \ldots)|^2 \, dx + \tag{1.5}$$

$$+ \int_{|(u(x), \ldots, D^\beta u(x), \ldots)| \leqq \delta} |g_j(x, u(x), \ldots, D^\beta u(x), \ldots)|^2 \, dx.$$

According to assumption (a_2) for any $\varepsilon > 0$ we can choose the number $\delta > 0$ such that

$$|g_j(x, u(x), \ldots, D^\beta u(x), \ldots)| \leqq \varepsilon \, |(u(x), \ldots, D^\beta u(x), \ldots)|.$$

Thus

$$\int_{|(u(x), \ldots, D^\beta (x), \ldots)| > \delta} |g_j(x, u(x), \ldots, D^\beta u(x), \ldots)|^2 \, dx \leqq$$

$$\leqq \varepsilon^2 \int_{B_a} |(u(x), \ldots, D^\beta u(x), \ldots,)|^2 \, dx \leqq \varepsilon^2 \, \|u\|^2_{\prod_{j=1}^{M} H^{2m}(B_a)}. \tag{1.6}$$

On r.h.s. of (1.5), the second term is bounded:

$$\int_{|(u(x), \ldots, D^\beta u(x), \ldots)| \leqq \delta} |g_j(x, u(x), \ldots, D^\beta u(x), \ldots)|^2 \, dx \leqq c, \tag{1.7}$$

where c is a constant, since g is continuous and $|x| \leqq a$, $|(u(x), \ldots, D^\beta u(x), \ldots)| \leqq \delta$. Therefore we obtain estimate (1.3) by the use of (1.6) and (1.7).

In order to prove that the equation (1.1) has at least one solution v belonging to product Hilbert space $\prod_{j=1}^{M} L_a^2(R^n)$ for any $f \in \prod_{j=1}^{M} L_a^2(R^n)$, we use Schauder's fixed point theorem. The estimation (1.3) implies that there may be chosen a number $\varrho_0 > 0$ such that

$$\|v\|_{\prod_{j=1}^{M} L_a^2(R^n)} > \varrho_0 \Rightarrow \frac{\|G(v)\|_{\prod_{j=1}^{M} L_a^2(R^n)}}{\|v\|_{\prod_{j=1}^{M} L_a^2(R^n)}} < \frac{1}{2}.$$

Let $F_j(v) = f_j - G_j(v)$, $f = (f_1, \ldots, f_M)$ and $F = (F_1, \ldots, F_M)$. Thus $F\colon \prod_{j=1}^{M} L_a^2(R^n) \to \prod_{j=1}^{M} H_a^1(R^n)$ is a bounded operator since $G\colon \prod_{j=1}^{M} L_a^2(R^n) \to \prod_{j=1}^{M} H_a^1(R^n)$ is a bounded operator and so

$$\|v\|_{\prod_{j=1}^{M} L_a^2(R^n)} \leqq \varrho_0 \Rightarrow \|F(v)\|_{\prod_{j=1}^{M} L_a^2(R^n)} \leqq \varrho_1.$$

Let ϱ be the number $\max\{\varrho_0, \varrho_1, 2\|f\|\}$. Then F maps the sphere $\{\in \prod_{j=1}^{M} L_a^2(R^n)$: $\|v\|_{\prod_{j=1}^{M} L_a^2(R^n)} \leqq \varrho\}$ into itself, as

$$\|F(v)\| \leqq \varrho_1 \leqq \varrho \quad \text{if} \quad \|v\| \leqq \varrho_0$$

and

$$\|F(v)\| \leqq \|f\| + \|G(v)\| \leqq \varrho \quad \text{if} \quad \varrho_0 \leqq \|v\| \leqq \varrho.$$

Also, F is a continuous, compact operator as is obvious from above. Hence, using Shauder's fixed point theorem, we come to the conclusion that F has at least one fixed point v belonging to $\prod_{j=1}^{M} L_a^2(R^n)$. This shows that the equation (1.1) has at least one solution. Thus $u = A^{-1}v$ will belong to $\prod_{j=1}^{M} H_{\text{loc}}^{2m}(R^n)$ and it will be a solution of the equation (0.1), which vanishes for $|x| \to \infty$.

We consider following boundary value problem over a ball

$$A_j u_j^\varrho + g_j(x, u^\varrho, \ldots, D^\beta u^\varrho, \ldots) = f_j \quad \text{in} \quad B_\varrho \tag{1.8}$$

$$B^{lj}(\omega, D)\, u_j^\varrho = 0 \quad \text{on} \quad S_\varrho, \tag{1.9}$$

$$j = 1, \ldots, M, \quad l = 1, \ldots, m.$$

THEOREM 1.2. *Let the conditions of Theorem 1.1 be satisfied and further assume that if* $\varrho \geqq \varrho_0$ *then for any* $f_j \in L_a^2(R^n)$ *the linear problem* (0.4) *has a unique solution* $u_j^\varrho \in H^{2m}(B_\varrho)$ *and the estimation* (0.7) *holds. Then for any* $\varrho \geqq \varrho_0$ *and* $f_j \in L_a^2(R^n)$ *the problem* (1.8), (1.9) *has at least one solution* $u^\varrho \in \prod_{j=1}^{M} H^{2m}(B_\varrho)$.

PROOF. Let us denote by $(A_{\varrho j})^{-1} f_j$ the unique solution $u_j^\varrho \in H^{2m}(B_\varrho)$ of the problem

$$A_j u_j^\varrho = f_j \quad \text{in} \quad B_\varrho$$

$$B^{lj}(x, D)\, u_j^\varrho = 0 \quad \text{on} \quad S_\varrho, \quad l = 1, \ldots, m,$$

and

$$A_\varrho^{-1} = (A_{\varrho 1}^{-1}, \ldots, A_{\varrho M}^{-1}).$$

If $v^\varrho = (v_1^\varrho, \ldots, v_M^\varrho)$ is a solution of

$$v_j^\varrho + g_j(x, A_\varrho^{-1} v^\varrho, \ldots, D^\beta A_\varrho^{-1} v^\varrho, \ldots) = f_j, \qquad j = 1, \ldots, M \tag{1.10}$$

then $u_j^\varrho = (A_{\varrho j})^{-1} v_j^\varrho \in H^{2m}(B_\varrho)$ is a solution of (1.8), (1.9). Let us define an operator G_ϱ by

$$G_{\varrho j}(v_j^\varrho) = g_j(x, A_\varrho^{-1} v^\varrho, \ldots, D^\beta A_\varrho^{-1} v^\varrho, \ldots),$$

and

$$G_\varrho = (G_{\varrho 1}, \ldots, G_{\varrho M}).$$

Then

$$G_\varrho\colon \prod_{j=1}^{M} L_a^2(R^n) \to \prod_{j=1}^{M} L_a^2(R^n)$$

is a compact operator as was shown previously in the proof of Theorem 1.1.

Further, as in the proof of Theorem 1.1, we can prove the equality

$$\lim_{\|v\|_{\prod_{j=1}^{M} L_a^2(R^n)} \to \infty} \frac{\|G_{\varrho j}(v)\|_{L_a^2(R^n)}}{\|v\|_{\prod_{j=1}^{M} L_a^2(R^n)}} = 0 \tag{1.11}$$

uniformly for $\varrho \geqq \varrho_0$. Applying Schauder's fixed point theorem, it can be proved that there exist solutions v_j^ϱ of equations (1.10). Hence $u_j^\varrho = (A_{\varrho j})^{-1} v_j^\varrho$ belongs to $H^{2m}(B_\varrho)$ and is a solution of (1.8), (1.9).

2. Approximation of solutions

THEOREM 2.1. *Let the conditions of Theorem 1.2 be fulfilled and consider a sequence of numbers* $\varrho_k \geqq \varrho_0$ *such that* $\lim \varrho_k = +\infty$. *If* u^{ϱ_k} *is a solution of* (1.8), (1.9) *for* $\varrho = \varrho_k$ *then* (ϱ_k) *has a subsequence* (ϱ_k^*) *such that*

$$\lim_{k \to \infty} \|u^{\varrho_k *} - u^*\|_{\prod_{j=1}^{M} H^{2m}(K)} = 0 \tag{2.1}$$

holds for any compact set $K \subset R^n$ *where* u^* *belongs to* $\prod_{j=1}^{M} H_{loc}^{2m}(R^n)$ *and it is a solution of* (0.1), *which vanishes for* $|x| \to \infty$.

If the solution u *of* (0.1) *is unique then for the solutions* u^{ϱ_k} *of* (1.8), (1.9) *the following holds for any* $K \subset R^n$:

$$\lim_{k \to \infty} \|u^{\varrho_k} - u\|_{\prod_{j=1}^{M} H^{2m}(K)} = 0. \tag{2.2}$$

PROOF. We show that the solutions u^{ϱ_k} are bounded in $\prod_{j=1}^{M} H^{2m}(B_a)$. Since $v^{\varrho_k} = (A_1 u_1^{\varrho_k}, \ldots, A_M u_M^{\varrho_k})$ satisfy (1.10) and estimation (1.11) holds thus v^{ϱ_k} are bounded in $\prod_{j=1}^{M} L_a^2(R^n)$.

Further, by (0.6), (0.7)

$$A^{-1}\colon \prod_{j=1}^{M} L_a^2(R^n) \to \prod_{j=1}^{M} H^{2m}(B_a)$$

is bounded linear operator and so the functions $u^\varrho = A_\varrho^{-1} v^\varrho$ are bounded in $\prod_{j=1}^{M} H^{2m}(B_a)$.

Therefore for all j there will be a subsequence $(u_j^{\varrho_k *})$ of $(u_j^{\varrho_k})$ which is convergent to an element belonging to $H^{2m-1}(B_a)$ in the norm of $H^{2m-1}(B_a)$:

$$\lim_{k \to \infty} \|u_j^{\varrho_k *} - u_j^0\|_{H^{2m-1}(B_a)} = 0. \tag{2.3}$$

According to assumption (a_3), the first derivatives are bounded and continuous. By the mean value theorem

$$|g_j(x, u_j^{\varrho_k *}, \ldots, D^\beta u_j^{\varrho_k *}, \ldots) - g_j(x, u^0, \ldots, D^\beta u^0, \ldots)| \leqq$$
$$\leqq c_1 \sum_{|\beta| \leqq 2m-1} |D^\beta u^{\varrho_k *} - D^\beta u^0|$$

where c_1 is a constant. Hence

(2.4) $$\lim_{k\to\infty} \int_{B_a} |g_j(x, u^{\varrho_k *}, \ldots, D^\beta u^{\varrho_k *}, \ldots) - g_j(x, u^0, \ldots, D^\beta u^0, \ldots)|^2 \, dx = 0.$$

Let us consider $v_j^{\varrho_k *} = A_j u_j^{\varrho_k *}$, then one may write

(2.5) $$v_j^{\varrho_k *} + g_j(x, u_j^{\varrho_k *}, \ldots, D^\beta u^{\varrho_k *}, \ldots) = f_j.$$

By the help of equations (2.4), (2.5) it can be proved that the sequence $(v_j^{\varrho_k *})$ tends to a function v_j^* belonging to $L_a^2(R^n)$ in the norm of $L_a^2(R^n)$. Thus we obtain

(2.6) $$v_j^* + g_j(x, u^0, \ldots, D^\beta u^0, \ldots) = f_j, \qquad j = 1, \ldots, M.$$

We show that for any compact $K \subset R^n$

(2.7) $$\lim_{k\to\infty} \|u_j^{\varrho_k *} - A_j^{-1} v_j^*\|_{H^{2m}(K)} = 0.$$

Since $u_j^{\varrho_k *} = (A_j^{\varrho_k *})^{-1} v^{\varrho_k *}$, thus

$$\|u_j^{\varrho_k *} - A_j^{-1} v_j^*\|_{H^{2m}(K)} \leqq \|(A_j^{\varrho_k *})^{-1} v_j^{\varrho_k *} - A_j^{-1} v_j^{\varrho_k *}\|_{H^{2m}(K)} + \|A_j^{-1}(v_j^{\varrho_k *} - v_j^*)\|_{H^{2m}(K)}.$$

Estimation (0.7) implies that the first term in the right of the above inequality tends to 0 as $k \to \infty$. The estimation (0.6) implies that $\|A_j^{-1}(v_j^{\varrho_k *} - v_j^*)\|_{H^{2m}(K)} \to 0$ as $k \to \infty$. Thus we proved

(2.8) $$\lim_{k\to\infty} \|u_j^{\varrho_k *} - A_j^{-1} v_j^*\|_{H^{2m}(K)} = 0.$$

From above it follows that

$$u_j^0 = A_j^{-1} v_j^* \quad \text{a. e. in the ball } B_a.$$

Let us put $A^{-1} v^* = u^*$, then $u_j^* = u_j^0$ a.e. in B_a, $v^* = A_j u_j^*$ and

(2.9) $$A_j u_j^* + g_j(x, u_j^*, \ldots, D^\beta u_j^*, \ldots) = f_j, \qquad j = 1, \ldots, M,$$

such that u^* tends to zero at infinity. Thus together with equations (2.9) and (2.7) we have proved that the estimation (2.1) is true.

The equality (2.2) can be proved by indirect method. Assume that the solution u of equations (0.1) is unique such that (2.2) does not hold. Then there exists a compact set K of R^n, a number $\varepsilon_0 > 0$ and also a subsequence $(u^{\tilde\varrho_k}) = (\tilde u^k)$ such that $\lim_{k\to\infty} \tilde\varrho_k = +\infty$, and

(2.10) $$\|\tilde u^k - u\|_{\prod_{j=1}^M H^{2m}(K)} \geqq \varepsilon_0.$$

The above result shows that there will be a subsequence $(\tilde{\tilde{u}}^k)$ of $(\tilde{u}^k)$ such that

$$\lim_{k\to\infty} \|\tilde{\tilde{u}}^k - \tilde{u}\|_{\prod_{j=1}^{M} H^{2m}(K)} = 0,$$

where $\tilde{u}$ is a solution of equation (0.1) which tends to zero at infinity. This means that $\tilde{u}=u$ because of the uniqueness of the solution but it is impossible because of (2.10).

REMARK 2.2. Assume that conditions of Theorem 2.1 are satisfied and also estimations (0.3), (0.8) hold, then

$$\lim_{k\to\infty} \|u^{\varrho_k *} - u^*\|_{\prod_{j=1}^{M} H^{2m}(B_{\varrho_k *})} = 0. \tag{2.11}$$

In the case, when the solution of (0.1) is unique, we have

$$\lim_{k\to\infty} \|u^{\varrho} - u\|_{\prod_{j=1}^{M} H^{2m}(B_{\varrho})} = 0. \tag{2.12}$$

The author acknowledges her gratefulness to Prof. László Simon for his valuable remarks and advices.

REFERENCES

[1] AGMON, S., DOUGLIS, A. and NIRENBERG, L., Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions. I, *Comm. Pure Appl. Math.* **12** (1959), 623—727. *MR* **23** # A2610.

[2] HÖRMANDER, L., *Linear partial differential operators,* Die Grundlehren der mathematischen Wissenschaften, Bd. 116, Academic Press, New York; Springer-Verlag, Berlin—Göttingen—Heidelberg, 1963. *MR* **28** # 4221.

[3] ШИМОН, Л., Об аппроксимации решений граничных задач в неограниченных областях, *Differencial'nye Uravnenija* **9** (1973), 1482—1492. *MR* **52** # 6178.

[4] SIMON, L., On elliptic differential equations in R^n, *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.* **27** (1985), 241—256.

[5] SIMON, L., On approximation of the solutions of quasi-linear elliptic equation in R^n, *Acta Sci. Math.* (*Szeged*) **47** (1984), 241—256.

(Received January 3, 1987)

EÖTVÖS LORÁND TUDOMÁNYEGYETEM
SZÁMÍTÓKÖZPONT
BOGDÁNFY U. 10/B.
H—1117 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 315—323

ON THE NON-HEREDITARINESS OF RADICAL AND SEMISIMPLE CLASSES OF NEAR-RINGS

S. VELDSMAN

Introduction

It is known that in the variety of near-rings, contrary to the associative ring case, not every semisimple class is hereditary. Betsch and Kaarli [3] has shown that no subidempotent radical can have a hereditary semisimple class. In the smaller variety of all 0-symmetric near-rings, most of the well-known radicals, e.g. J_2, J_3 and the Brown—McCoy radical have hereditary semisimple classes. In fact, they satisfy a stronger condition, namely they are ideal-hereditary, c.f. Kaarli [9], Holcombe [8] and Anderson, Kaarli and Wiegandt [1]. The nilradical is hereditary, but Kaarli [10] has shown that its semisimple class is not hereditary. We will show that in the variety of all, not necessarily 0-symmetric near-rings, "good" radicals which are ideal-hereditary seem to be rare: every radical for which the two element field is semisimple cannot be ideal-hereditary. Amongst others, we also show that J_2 is not a Kurosh—Amitsur radical in the class of all near-rings.

1. Preliminaries

Near-ring notions will be from Pilz [14], hence all near-rings will be right near-rings. For a near-ring N, N^+ will denote the underlying group of N and N^c will be the constant near-ring built on N^+ by the multiplication $ab=a$ for all $a, b\in N$. The two element field will be denoted by I_2. By Z we will denote the class of all zero-near-rings, i.e. $Z=\{N|N^2=0\}$. As usual, $I\lhd N$ means that I is an ideal in N.

We have to fix some notation and recall a few definitions and results. For Kurosh—Amitsur radicals, the basics can be found in Wiegandt [18] and for Hoehnke radicals we refer to Hoehnke [7].

1.1 Let $\mathcal{M}$ be a class of near-rings. Classes of near-rings will always be assumed to be abstract, i.e. they contain the one-element near-ring and are closed under isomorphic copies. With every near-ring N we associate two ideals of N, depending on $\mathcal{M}$, defined by

$$\mathcal{M}(N) := \sum (I\lhd N | I\in\mathcal{M})$$

and

$$(N)\mathcal{M} := \bigcap\{I\lhd N | N/I\in\mathcal{M}\}.$$

As usual, $\mathcal{M}$ is *regular* if $(I)\mathcal{M}\neq I$ for each $0\neq I\lhd N\in\mathcal{M}$ and *hereditary* if $I\lhd N\in\mathcal{M}$ implies $I\in\mathcal{M}$.

1980 *Mathematics Subject Classification* (1985 Revision). Primary 16A76; Secondary 16A21.
Key words and phrases. Hoehnke radical, Kurosh—Amitsur radical, hereditariness.

1.2 Let ϱ be a mapping which assigns to each near-ring N an ideal ϱN of N. Such mappings will be called *ideal-mappings*. We will consider the following conditions that ϱ may satisfy:
(H1) $g(\varrho N)\subseteq\varrho(g(N))$ for all homomorphisms $g\colon N\to N'$;
(H2) $\varrho(N/\varrho N)=0$ for all N;
r-hereditary if $I\cap\varrho N\subseteq\varrho I$ for all $I\lhd N$;
s-hereditary if $\varrho I\subseteq I\cap\varrho N$ for all $I\lhd N$;
ideal-hereditary if it is both r-hereditary and s-hereditary, i.e. if $\varrho I=I\cap\varrho N$ for all $I\lhd N$;
idempotent if $\varrho\varrho N=\varrho$ for all N;
complete if $\varrho I=I\lhd N$ implies $I\subseteq\varrho N$;
right strong if I is a right invariant subgroup of N and $\varrho I=I$, then $I\subseteq\varrho N$;
hypersolvable if $N\in Z$ implies $\varrho N=N$;
ADS-property if $\varrho I\lhd N$ for all $I\lhd N$;
With ϱ we associate two classes of near-rings $\mathscr{R}_\varrho$ and $\mathscr{S}_\varrho$ defined by
$\mathscr{R}_\varrho:=\{N|\varrho N=N\}$ and
$\mathscr{S}_\varrho:=\{N|\varrho N=0\}$ which will be called the ϱ-radical class and ϱ-semisimple class, respectively.

Note that $\mathscr{R}_\varrho\cap\mathscr{S}_\varrho=0$ and that every simple near-ring is in $\mathscr{R}_\varrho$ or in $\mathscr{S}_\varrho$. Furthermore, ϱ is s-hereditary iff the class $\mathscr{S}_\varrho$ is hereditary. If ϱ is r-hereditary, then the class $\mathscr{R}_\varrho$ is hereditary, the converse is not true in general. The next result is well known and the proof is straightforward:

PROPOSITION 1.2.1. *Let ϱ be an ideal-mapping. Then*
(i) *ϱ r-hereditary implies ϱ idempotent and*
(ii) *ϱ s-hereditary implies ϱ complete.*

If ϱ and ϱ' are two ideal-mappings such that $\varrho N\subseteq\varrho' N$ for all N, we will denote this by $\varrho\leqq\varrho'$. In such a case, $\mathscr{R}_\varrho\subseteq\mathscr{R}_{\varrho'}$ and $\mathscr{S}_{\varrho'}\subseteq\mathscr{S}_\varrho$.

1.3. An ideal-mapping ϱ is a *Hoehnke radical* (H-radical) if it satisfies conditions (H1) and (H2).

In such a case, the class $\mathscr{R}_\varrho$ is always homomorphically closed as can easily be verified by using condition (H1).

Let $\mathscr{M}$ be a class of near-rings. Then ϱ defined by $\varrho N:=(N)\mathscr{M}$ is always an H-radical with $\mathscr{M}\subseteq\mathscr{S}_\varrho$. In fact, $\mathscr{S}_\varrho$ is the subdirect closure of $\mathscr{M}$. Mlitz [13] has shown that every H-radical can be obtained as above for a suitable class $\mathscr{M}$.

Note that $\mathscr{M}$ is regular iff ϱ is complete and that if N is a simple near-ring, then $N\in\mathscr{M}$ iff $N\in\mathscr{S}_\varrho$.

1.4. An ideal-mapping ϱ is a *Kurosh—Amitsur radical* (KA-radical) if it is a complete idempotent H-radical.

A class of near-rings $\mathscr{R}$ is a *Kurosh—Amitsur radical class* if it satisfies the following conditions:

(R1) $\mathscr{R}$ is homomorphically closed.

(R2) $\mathscr{R}(N)\in\mathscr{R}$ for all N.

(R3) $\mathscr{R}(N/\mathscr{R}(N))=0$ for all N.

Mlitz [12] has shown that there is a one-to-one correspondence between KA-radical classes and KA-radical mappings: If $\mathcal{R}$ is a KA-radical class, then ϱ defined by $\varrho N := \mathcal{R}(N)$ is a KA-radical mapping with $\mathcal{R}_\varrho = \mathcal{R}$. Conversely, if ϱ is a KA-radical mapping, then $\mathcal{R}_\varrho$ is a KA-radical class with $\varrho N = \mathcal{R}_\varrho(N)$.

Any regular class $\mathcal{M}$ determines a KA-radical class, namely the *upper radical class* $\mathcal{U}\mathcal{M}$ which is defined by

$$\mathcal{U}\mathcal{M} := \{N|(N)\mathcal{M} = N\}.$$

If $\mathcal{R}$ is a KA-radical class, we usually denote its semisimple class by $\mathcal{S}\mathcal{R} := \{N|\mathcal{R}(N)=0\}$. It is well-known that $\mathcal{R}(N)=(N)\mathcal{S}\mathcal{R}$ for all N. Furthermore, $\mathcal{R}$ is hereditary iff $I \cap \mathcal{R}(N) \subseteq \mathcal{R}(I)$ for all $I \lhd N$. Any ideal-hereditary H-radical is a KA-radical mapping (cf. Proposition 1.2.1).

Let $\mathcal{M}$ be a regular class of near-rings. Let ϱ be the H-radical defined by $\varrho N = (N)\mathcal{M}$. Let $\mathcal{R}$ be the upper radical class determined by $\mathcal{M}$, i.e. $\mathcal{R} := \mathcal{U}\mathcal{M}$. Then $\mathcal{R} = \mathcal{R}_\varrho$, $\mathcal{S}_\varrho = \bar{\mathcal{M}} \subseteq \mathcal{S}\mathcal{R}$ (where $\bar{\mathcal{M}}$ is the subdirect closure of $\mathcal{M}$) and $\mathcal{R}(N) \subseteq \varrho N$ for all N. Although, in this case, $\mathcal{R}_\varrho$ is a KA-radical class, ϱ need not be a KA-radical mapping. If $\mathcal{R}(N) = \varrho N$ for all N, which is equivalent to assuming $\bar{\mathcal{M}} = \mathcal{S}\mathcal{R}$, then ϱ is a KA-radical mapping.

We will need the following result due to Betsch and Kaarli:

THEOREM **1.5** [3]. *If* $\mathcal{R} \neq 0$ *is a KA-radical class such that* $\mathcal{S}\mathcal{R}$ *is hereditary, then* $Z \subseteq \mathcal{R}$.

1.6 Our results are based on Example no. 11 on the dihedral group D_8 given in Pilz [14], p. 345.

For ease of reference, we give both the multiplication and addition tables (in a different notation):

+	0	1	2	3	4	5	6	7
$0=0$	0	1	2	3	4	5	6	7
$a=1$	1	2	3	0	5	6	7	4
$2a=2$	2	3	0	1	6	7	4	5
$3a=3$	3	0	1	2	7	4	5	6
$b=4$	4	7	6	5	0	3	2	1
$a+b=5$	5	4	7	6	1	0	3	2
$2a+b=6$	6	5	4	7	2	1	0	3
$3a+b=7$	7	6	5	4	3	2	1	0

·	0	1	2	3	4	5	6	7
0	0	0	0	0	0	0	0	0
1	0	1	0	1	0	1	1	0
2	0	2	0	2	0	2	2	0
3	0	3	0	3	0	3	3	0
4	4	4	4	4	4	4	4	4
5	4	5	4	5	4	5	5	4
6	4	6	4	6	4	6	6	4
7	4	7	4	7	4	7	7	4

In the sequel, this non 0-symmetric near-ring will be denoted by W. Let $V=\{0, 1, 2, 3\}$, $U=\{0, 2\}$, $X=\{0, 2, 4, 6\}$ and $Y=\{0, 4\}$.

Certain substructures of W are:

Ideals of W: O, V, W

Right ideals of W: O, U, V, X, W

Left ideals of W: O, V, W

Ideals of V: O, U, V

Left ideals of V: O, U, V

Right ideals of V: O, U, V

Note that $W^2=W$, $V^2=V$, $U^2=O$, $W/V\cong I_2^c$ and $V/U\cong I_2$.

1.7 Although we will work in the variety of near-rings, our results remain true also in the class of all finite near-rings. In fact, they are valid in any universal class of near-rings which contains W and in which Theorem 1.5 remains true. The proof of this theorem depends on the existence of two constructions of near-rings in the universal class. A sufficient condition therefore is that for any two near-rings N_1 and N_2 in the universal class, any near-ring on $N_1^+\oplus N_2^+$ must also be in the universal class.

2. Non-hereditary semisimple classes

Let $\mathcal{M}$ be a fixed class of semiprime near-rings. In this section, ϱ will be the ideal mapping defined by $\varrho N:=(N)\mathcal{M}$ for all N. This means that ϱ is a hypersolvable H-radical.

THEOREM 2.1. (i) *If* $W\in\mathcal{S}_\varrho$, *then* ϱ *is not s-hereditary.* (ii) *If* I_2 *and* I_2^c *are in* $\mathcal{M}$ *and* ϱ *is idempotent, then* ϱ *is not s-hereditary.*

PROOF. (i) If ϱ is s-hereditary, then $U\lhd V\lhd W\in\mathcal{S}_\varrho$ implies $U\in\mathcal{S}_\varrho\cap Z\subseteq\mathcal{S}_\varrho\cap\cap\mathcal{R}_\varrho=0$; a contradiction.

(ii) If ϱ is idempotent and I_2 and I_2^c are in $\mathcal{M}$, then $W\in\mathcal{S}_\varrho$. Indeed, the only other possibilities are $\varrho W=W$ or $\varrho W=V$. But neither of these two are possible because $\mathcal{R}_\varrho$ is homomorphically closed, ϱ is idempotent and I_2 and I_2^c are in $\mathcal{M}\subseteq\mathcal{S}_\varrho$. Hence $W\in\mathcal{S}_\varrho$ and the result follows from (i) above.

COROLLARY 2.2. *Let* $\mathcal{R}\neq 0$ *be a KA-radical class such that both* I_2 *and* I_2^c *are in* $\mathcal{S}\mathcal{R}$. *Then* $\mathcal{S}\mathcal{R}$ *is not hereditary.*

PROOF. If $\mathcal{S}\mathcal{R}$ is hereditary, then $Z\subseteq\mathcal{R}$ (by Theorem 1.5) which means that $\mathcal{S}\mathcal{R}$ must consist of semiprime near-rings. Because $\mathcal{R}$ is a KA-radical class, $\mathcal{R}(N)=(N)\mathcal{S}\mathcal{R}$ and $\mathcal{R}(\mathcal{R}(N))=\mathcal{R}(N)$ for all N. The assertion then follows from the theorem.

THEOREM 2.3. *If* $I_2\in\mathcal{M}$, *then* ϱ *is not ideal-hereditary.*

PROOF. If ϱ is ideal-hereditary, then $\varrho V = V \cap \varrho W \lhd W$. But $V/U \cong I_2 \in \mathcal{M}$ implies $\varrho V \subseteq U$. Hence $\varrho V = 0$. Since $\mathcal{S}_\varrho$ is hereditary, $U \in \mathcal{S}_\varrho \cap Z \subset \mathcal{S}_\varrho \cap \mathcal{R}_\varrho = 0$, which is not possible.

COROLLARY 2.4. *If $\mathcal{R} \neq 0$ is a KA-radical class such that $I_2 \in \mathcal{S}\mathcal{R}$, then $\mathcal{R}$ does not satisfy the ADS-property and is hence not ideal-hereditary.*

The proof is as in the proof of the above corollary. From Theorem 2.3 it is also clear that if ϱ is r-hereditary and $I_2 \in \mathcal{M}$, then ϱ is not s-hereditary. More criteria for the non-hereditariness of $\mathcal{S}_\varrho$, using W again, is given by:

PROPOSITION 2.5. *Suppose ϱ is idempotent and satisfy any one of the following three conditions:*

(1) *$W \notin \mathcal{R}_\varrho$ and $I_2 \in \mathcal{M}$.*

(2) *$\mathcal{R}_\varrho$ contains no idempotent near-rings.*

(2) *$\mathcal{R}_\varrho$ contains no near-rings which have right identities. Then ϱ is not s-hereditary.*

PROOF. Suppose ϱ is s-hereditary.

(1) If $W \notin \mathcal{R}_\varrho$ and $I_2 \in \mathcal{M}$, then $\varrho W = V$. The idempotence of ϱ implies that $V \in \mathcal{R}_\varrho$. Because $\mathcal{R}_\varrho$ is homomorphically closed, we have $V/U \cong I_2 \in \mathcal{R}_\varrho$ which contradicts $I_2 \in \mathcal{M}$.

(2) In this case $\varrho W = 0$ because both W and V are idempotent near-rings. The result then follows from Theorem 2.1.

(3) As in (2) above because both W and V has right identities.

COROLLARY 2.6. *Let $\mathcal{R} \neq 0$ be a KA-radical class. Suppose $\mathcal{R}$ satisfies any one of the following three conditions:*

(1) *$W \notin \mathcal{R}$ and $I_2 \in \mathcal{S}\mathcal{R}$.*

(2) *$\mathcal{R}$ contains no idempotent near-rings.*

(3) *$\mathcal{R}$ contains no near-rings with right identities.*

Then $\mathcal{S}\mathcal{R}$ is not hereditary.

3. Non-hereditary radical classes

PROPOSITION 3.1. *Let ϱ be any ideal-mapping which satisfies condition* (H1). *If $I_2 \in \mathcal{S}_\varrho$ and $W \in \mathcal{R}_\varrho$, then $\mathcal{R}_\varrho$ is not hereditary, hence ϱ is not r-hereditary.*

PROOF. If $\mathcal{R}_\varrho$ is hereditary, then $V \lhd W \in \mathcal{R}_\varrho$ implies $V \in \mathcal{R}_\varrho$. But condition (H1) then yields $I_2 \in \mathcal{R}_\varrho$ which contradicts $I_2 \in \mathcal{S}_\varrho$.

COROLLARY 3.2. *Any KA-radical class for which $W \in \mathcal{R}$ and $I_2 \in \mathcal{S}\mathcal{R}$ is not hereditary.*

In view of Proposition 1.2.1, conditions yielding ϱ not idempotent also implies the non-(r-hereditariness) of ϱ. Such a mapping is of course not a KA-radical mapping.

PROPOSITION 3.3. *Let ϱ be an ideal-mapping which satisfies condition* (H1) *and for which $\varrho W = V$ and $I_2 \in \mathcal{S}_\varrho$. Then ϱ is not idempotent and hence not a KA-radical mapping.*

PROOF. If ϱ is idempotent, then $V\in\mathcal{R}_\varrho$ which yields $I_2\in\mathcal{R}_\varrho$ by condition (H1). This contradicts $I_2\in\mathcal{S}_\varrho$.

We conclude this section with a result on right strong radicals which has recently been studied by Anderson, Kaarli and Wiegandt [2] (in fact, left strong radicals because they deal with left near-rings).

PROPOSITION 3.4. *Let ϱ be an ideal-mapping of the form $\varrho N:=(N)\mathcal{M}$ for some class of semiprime near-rings $\mathcal{M}$. If $W\in\mathcal{S}_\varrho$, then ϱ is not right strong.*

PROOF. Because U is a right invariant subgroup of W and $U^2=0$, we would have $U\subseteq\varrho W=0$ if ϱ is right strong.

COROLLARY 3.5. *Let $\mathcal{R}\neq 0$ be a KA-radical class. If $W\in\mathcal{S}\mathcal{R}$, then $\mathcal{R}$ is not right strong.*

PROOF. If $\mathcal{R}$ is right strong, then Theorem 1 in Anderson, Kaarli and Wiegandt [2] implies that $\mathcal{R}$ is hypersolvable. Hence $\mathcal{S}\mathcal{R}$ consists of semiprime near-rings and because $\mathrm{R}(N)=(N)\mathcal{S}\mathrm{R}$ for all N, the result follows from the above proposition.

4. Applications

4.1 Let $\mathcal{N}$ be the nil radical class, i.e. $\mathcal{N}=\{N\mid N$ is a nil near-ring$\}$. It is well-known that $\mathcal{N}$ is a hereditary KA-radical class. Earlier, Kaarli [10] has given an example of a 0-symmetric near-ring which shows that $\mathcal{S}\mathcal{N}$ is not hereditary (hence not ideal-hereditary and it does not satisfy the ADS-property). Anderson, Kaarli and Wiegandt [2] has also pointed out that this example shows that $\mathcal{N}$ is not right strong. The near-ring W also demonstrates all these (negative) properties.

Due to the lack of symmetry for near-rings, one can also ask if $\mathcal{N}$ is left strong in the variety of all (right) near-rings. This is not the case, as example we use the mentioned example of Kaarli: Let G be a finite abelian group which contains a proper subgroup H. Let $N:=\{f\in M_0(G)\mid f(H)=0\}$ and let $L=(H:G)_N:=\{f\in N\mid f(G)\subseteq H\}$. Then L is a non-zero left invariant subgroup of N, $L^2=0$ and because N is simple, it can easily be verified that $N\in\mathcal{S}\mathcal{N}$.

4.2 Let $\mathcal{L}$ be the class of all locally nilpotent near-rings. We recall, a near-ring N is *locally nilpotent* if every finite subset of N is nilpotent. Bhandari and Saxena [4] have shown that $\mathcal{L}$ is a hereditary KA-radical class. Since I_2 and W are in $\mathcal{S}\mathcal{L}$, $\mathcal{L}$ does not satisfy the ADS-property, is not ideal-hereditary nor right strong and $\mathcal{S}\mathcal{L}$ is not hereditary. These conclusions contradict Theorem 1.10 in [6].

4.3 Let β be the lower Baer radical, i.e. $\beta=\mathcal{L}\mathcal{A}$ where $\mathcal{A}$ is the class of all nilpotent near-rings and $\mathcal{L}$ is the lower radical operator (cf. Tangeman and Kreiling [15]). Because the hereditariness of $\mathcal{A}$ is retained under the lower radical construction, β is hereditary. Since I_2 and W are in $\mathcal{S}\beta$, β does not satisfy the ADS-property, is not ideal-hereditary, is not right strong and $\mathcal{S}\beta$ is not hereditary.

4.4 Let $\mathcal{G}$ be the (KA) Brown—McCoy radical class, i.e. $\mathcal{G}$ is the upper radical determined by the class of simple near-rings with identity. Since $W\in\mathcal{G}$ and $I_2\in\mathcal{S}\mathcal{G}$,

$\mathcal{G}$ is not hereditary, does not satisfy the ADS-property and is not ideal-hereditary. $\mathcal{G}$ is not right strong (not even in the associative ring case). It is not known if $\mathcal{S}\mathcal{G}$ is hereditary or not.

4.5 Let $\mathcal{B}$ be any class of near-fields which contains I_2 and I_2^c. Let $\mathcal{F}:=\mathcal{U}\mathcal{B}$. Then $W\in\mathcal{S}\mathcal{F}$ and because I_2 and I_2^c are in $\mathcal{S}\mathcal{F}$, $\mathcal{F}$ is not ideal-hereditary, does not satisfy the ADS-property, is not right strong and $\mathcal{S}\mathcal{F}$ is not hereditary.

4.6 Let ϱ be the prime radical, i.e. $\varrho N=(N)\mathcal{M}$ where $\mathcal{M}$ is the class of all prime near-rings. Because I_2, I_2^c and W are in $\mathcal{M}$, ϱ is not s-hereditary and not right strong. It is not known if ϱ is r-hereditary or whether ϱ is a KA-radical mapping.[1]

4.7 A near-ring N is *s-prime* if $N\setminus\{0\}$ contains a multiplicative closed subset S such that $(a)\cap S\neq\varphi$ for all $0\neq a\in N$ where (a) is the ideal in N generated by a (cf. Van der Walt [16]). Let $\mathcal{M}$ be the class of all s-prime near-rings and let ϱ be the Hoehnke radical defined by $\varrho N:=(N)\mathcal{M}$. Since I_2, I_2^c and W are in $\mathcal{M}$, ϱ is not s-hereditary. We can also mention the following which is due to Groenewald [5]:

A subset U of a near-ring N is complete if $x\in U$ implies $x^n\in U$ where n is any positive integer. A near-ring N is *s-semiprime* if $N\setminus\{0\}$ contains a complete subset U such that $(a)\cap U\neq\emptyset$ for all $0\neq a\in N$. Let $\varrho'N:=(N)\mathcal{M}'$ where $\mathcal{M}'$ is the class of all s-semiprime near-rings. Then $\varrho'N=\varrho N$ for all N. This means ϱ' is not s-hereditary which points to an error in [5].

4.8 Let $\mathcal{M}$ be the class of all completely prime near-rings (i.e. near-rings N such that $ab\neq 0$ for all $a\neq 0$, $b\neq 0$ in N). Let $\varrho N:=(N)\mathcal{M}$. Then I_2 and I_2^c are in $\mathcal{M}$ and $\varrho W=V$. This means ϱ is not idempotent and hence not a KA-radical mapping and neither r-hereditary. It can easily be verified that ϱ is s-hereditary.

4.9 For each $\nu\in\{0, 1, 2, 3\}$, let $\mathcal{M}_\nu$ be the class of ν-primitive near-rings. We might recall the definition of a 3-primitive near-ring. A near-ring N is *3-primitive* if there exists a faithful N-group of type 3 where an N-group G is of type 3 if it is of type 2 and satisfies: $ng_1=ng_2$ for all $n\in N$ implies $g_1=g_2$ where $g_1, g_2\in G$ (cf. Holcombe [8]).

It is well-known that $\mathcal{M}_3\subseteq\mathcal{M}_2\subseteq\mathcal{M}_1\subseteq\mathcal{M}_0$ and consequently that $\mathcal{J}_0\leqq\mathcal{J}_1\leqq\mathcal{J}_2\leqq\leqq\mathcal{J}_3$ where $\mathcal{J}_\nu N:=(N)\mathcal{M}_\nu$ for each ν. Kaarli [11] has given examples of 0-symmetric near-rings to show that $\mathcal{J}_0$ and $\mathcal{J}_1$ are not idempotent; hence they are not KA-radicals and not r-hereditary. It is known that $\mathcal{J}_1$ is s-hereditary, but $\mathcal{J}_0$ is not s-hereditary (cf. Pilz [14], 5.19(b) for an example). In the variety of 0-symmetric near-rings, both $\mathcal{J}_2$ and $\mathcal{J}_3$ are ideal-hereditary KA-radicals. However, in the variety of all near-rings, $\mathcal{J}_2$ is not a KA-radical for it is not idempotent. It can easily be verified that $\mathcal{J}_0W=\mathcal{J}_1W=\mathcal{J}_2W=V$ and that both I_2 and I_2^c are in $\mathcal{M}_2$. By Proposition 3.3 we conclude that $\mathcal{J}_\nu$ is not idempotent, hence not KA-radical or r-hereditary, for $\nu=0, 1, 2$. Furthermore, since $\mathcal{J}_3W=W$ and $I_2\in\mathcal{M}_3$, $\mathcal{J}_3$ is not r-hereditary. It is not known whether $\mathcal{J}_3$ is a KA-radical in the variety of all near-rings.

---

[1] ADDED in text: K. Kaarli and T. Kriis (On the prime radical of near-rings, *Tartu Riikl. Ül. Toimetised* **764** (1987), 23-29) have shown that, for 0-symmetric near-rings, $\mathcal{L}$ is Γ-hereditary but not a KA-radical.

4.10 We conclude with two general remarks:

(1) If $\mathcal{M}$ is any class of semiprime near-rings which contains W, then $\mathcal{M}$ is not hereditary. Thus the classes of all semiprime near-rings, prime near-rings or s-prime near-rings are not hereditary.

(2) A condition that is often used in the study of supernilpotent radicals is the following:

A class of near-rings $\mathcal{M}$ satisfies condition (F) if $J \lhd K \lhd N$ and $K/J \in \mathcal{M}$ implies $J \lhd N$.

Sometimes the weaker condition (G) would suffice:

(G) If $J \lhd K \lhd N$ with $K/J \in \mathcal{M}$, then there exists an ideal $L \lhd N$ such that $L \subseteq J$ and $K/L \in \mathcal{M}$.

Any class of semiprime associative or alternative rings satisfies condition (F). Anderson, Kaarli and Wiegandt [1] has shown that in quite a few varieties of Ω-groups, e.g. those of associative rings, alternative rings, near-rings or 0-symmetric near-rings, any regular class of Ω-groups which satisfy condition (F) must consist of semiprime Ω-groups. Although every semisimple class in the varieties of associative or alternative rings must satisfy condition (G), it was shown in [17] that any regular class of near-rings which satisfies condition (G) must consist of semiprime near-rings. The converse is not true: Let $\mathcal{M}$ be a regular class of semiprime near-rings which contains I_2. Then $\mathcal{M}$ does not satisfy condition (G), and hence not (F), as can easily be seen from $U \lhd V \lhd W$ with $V/U \cong I_2 \in \mathcal{M}$.

REFERENCES

[1] Anderson, T., Kaarli, K. and Wiegandt, R., Radicals and subdirect decompositions, *Comm in Algebra* **13** (1985), 479—494.

[2] Anderson, T., Kaarli, K. and Wiegandt, R., On left strong radicals of near-rings *Proc. Edinburgh Math. Soc.* **31** (1988), 447—456.

[3] Betsch, G. and Kaarli, K., Supernilpotent radicals and hereditariness of semisimple classes of near-rings, *Radical Theory*, Coll. Math. Soc. J. Bolyai, 38, Eger, 1982, North-Holland, 1985, 47—58.

[4] Bhandari, M. C. and Saxena, P. K., A note on Levitzki radical of near-rings, *Kyungpook Math. J.* **20** (1980), 183—188. *MR* **82i**: 16038.

[5] Groenewald, N. J., Strongly semiprime ideals in near-rings, *Chinese J. Math.* **11** (1983), 221—227. *MR* **85d**: 16030.

[6] Groenewald, N. J. and Potgieter, P. C., A note on the Levitzki radical of a near-ring, *J. Austral. Math. Soc. Ser. A* **36** (1984), 416—420. *MR* **85j**: 16057.

[7] Hoehnke, H.-J., Radikale in allgemeinen Algebren, *Math. Nachr.* **32** (1966), 347—383. *MR* **37** #2662.

[8] Holcombe, M., A hereditary radical for near-rings, *Studia Sci. Math. Hungar.* **17** (1982), 453—456. *MR* **85i**: 16049.

[9] Kaarli, K., Radicals in near-rings, *Tartu Riikl. Ül. Toimetised Vih.* **390** (1976), 134—171 (in Russian). *MR* **57** #9760.

[10] Kaarli, K., Classification of irreducible R-groups over a semiprimary near-ring, *Tartu Riikl. Ül. Toimetised,* no. 556 (1981), 47—63 (in Russian). *MR* **82k**: 16047.

[11] Kaarli, K., On Jacobson type radicals of near-rings, *Acta. Math. Hungar.* **50** (1987), 71—78.

[12] Mlitz, R., Radicals and semisimple classes of Ω-groups, *Proc. Edinburgh Math. Soc.* (2) **23** (1980), 37—41. *MR* **82e**: 17005.

[13] Mlitz, R., Radicals and interpolation in universal algebra, *Radical Theory,* Coll. Math. Soc. J. Bolyai, 38, Eger, 1982, North-Holland, 1985, 291—331.

[14] Pilz, G., *Near-rings. The theory and its applications*, North-Holland Mathematics Studies, no. 23, North-Holland Publishing Co., Amsterdam—New York—Oxford, 1977. *MR* **57** #9761.
[15] Tangeman, R. and Kreiling, D., Lower radicals in nonassociative rings, *J. Austral. Math. Soc.* **14** (1972), 419—423. *MR* **47** #3470.
[16] Van der Walt, A. P. J., Prime ideals and nil radicals in nearrings, *Arch. Math. (Basel)* **15** (1964), 408—414.
[17] Veldsman, S., Supernilpotent radicals of near-rings, *Comm. in Algebra* **15** (1987), 2497-2510.
[18] Wiegandt, R., *Radical and semisimple classes of rings*, Queen's Papers in Pure and Applied Mathematics, No. 37, Queen's University, Kingston, Ontario, 1974. *MR* **50** #2227.

(Received February 3, 1987)

UNIVERSITEIT VAN PORT ELIZABETH
POSBUS 1600
ZA—6000 PORT ELIZABETH
REPUBLIC OF SOUTH AFRICA

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 325—329

ON A POLYNOMIAL ALGORITHM FOR SELECTING A LATTICE-BASIS CONTAINING A GIVEN PRIMITIVE SYSTEM

ÁKOS G. HORVÁTH

Introduction

It was proved by van der Waerden [1] that a vector system having k elements in an n-dimensional lattice $(k \leqq n)$ is a primitive system if and only if it can be supplemented for a basis of the lattice. By the help of a new, constructive proof of this theorem that we give now we can supplement a given primitive system for a basis of the lattice at polynomial speed $(O(n^3))$. This algorithm can play an important role in solving the reductional problem.

Definitions, notations

E^n: Euclidean n-space;

Z: the set of integers;

n-lattice: Let $\mathbf{r}_1, \dots, \mathbf{r}_n$ be linearly independent vectors in E^n. Let $[\mathbf{r}_1, \dots, \mathbf{r}_n] = \{\mathbf{x} \in E^n \ \mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{r}_i \ \forall i \ \ 1 \leqq i \leqq n, \ \ x_i \in Z\}$ $[\mathbf{r}_1, \dots, \mathbf{r}_n]$ is called n-lattice with basis $\{\mathbf{r}_i\}$, $i=1, \dots, n$;

primitive system: Let Γ be an n-lattice and let $\mathbf{a}_1, \dots, \mathbf{a}_k$ $(1 \leqq k \leqq n)$ be linearly independent vectors of Γ. We say that $\{\mathbf{a}_i\}$, $i=1, \dots, k$ is a primitive system in Γ if it is a basis of the sublattice lying in their subspace. Denote by $L[\mathbf{a}_1, \dots, \mathbf{a}_k]$ the subspace which is spanned by $\{\mathbf{a}_i\}$, $i=1, \dots, k$. Obviously, $\{\mathbf{a}_i\}$ $i=1, \dots, k$ is primitive system in Γ iff the following condition holds: $[\mathbf{a}_1, \dots, \mathbf{a}_k] = \Gamma \cap L[\mathbf{a}_1, \dots, \mathbf{a}_k]$;

primitive vector: primitive system consisting of one element;

g.c.d. $(x_1, \dots, x_n)$ is the greatest common divisor of the integer numbers $x_1, \dots, x_n$.

Characterizations of primitive systems

THEOREM 1 (van der Waerden) [1]. *Let Γ be an n-dimensional lattice and let $\{\mathbf{p}_1, \dots, \mathbf{p}_l\}$ be linearly independent vector system of this lattice. The system $\{\mathbf{p}_1, \dots, \mathbf{p}_l\}$ is a primitive one iff it can be supplemented for a basis of Γ (viz. there exists a $\{\mathbf{p}_1, \dots, \mathbf{p}_n\}$ vector system so that $\Gamma = [\mathbf{p}_1, \dots, \mathbf{p}_n]$).*

1980 *Mathematics Subject Classification* (1985 Revision). Primary 10E25; Secondary 10E05.
Key words and phrases. Lattice, base, primitive system.

For the proof of the theorem the algebraic characterizations of primitive systems will be used.

Theorem 2 (G. Horváth Ákos, 1985) [2]. *Let $\Gamma := [f_1, \dots, f_n]$ be an n-dimensional lattice and let $\mathbf{p}_1, \dots, \mathbf{p}_l$ be linearly independent vectors of Γ. Denote by A the matrix of the coefficients of $\{\mathbf{p}_i\}$ $(i=1, \dots, l)$ $(A \in \mathbf{Z}^{n\times l})$ and let* $\det \tilde{A}, \dots, \det \overset{\binom{n}{k}}{\tilde{A}}$ *be the l-dimensional subdeterminants of A. Then the system $\{\mathbf{p}_1, \dots, \mathbf{p}^l\}$ is a primitive one iff following condition holds:*

$$\text{g.c.d.} \det \left(\tilde{A}, \dots, \det \overset{\binom{n}{l}}{\tilde{A}}\right) = 1. \tag{1}$$

Lemma 1. *Let $A \in \mathbf{Z}^{n\times k}$ and $c \in \mathbf{Z}$ be given and let d be the greatest common divisor (g.c.d.) of the determinants of submatrices of A which take $l \times l$ shape. Assume that from one of the rows of A we subtract an other row which is multiplied by c. Then the greatest common divisor of the determinants of the submatrices (taking $l \times l$ shape) of the new $n \times l$-form matrix we have got is also d.*

Proof. Denote by $\mathbf{a}_{1_i}, \dots, \mathbf{a}_{l_i}$ the elements of i-th row of A in order and let $\tilde{A} \dots \overset{\binom{n}{l}}{\tilde{A}}$ be those submatrices of A which take $l \times l$-shape. Let B be written in the following form:

$$B := \begin{pmatrix} a_{11} & \dots a_{l_1} \\ \vdots & \vdots \\ a_{1_i} & a_{l_i} \\ \vdots & \vdots \\ a_{1_j} - c a_{1_i} & \dots a_{l_j} - c a_{l_i} \\ \vdots & \vdots \\ a_{1_n} & a_{l_n} \end{pmatrix} \in \mathbf{Z}^{n \times l}.$$

Let us denote by $\tilde{B}$ that $l \times l$-form submatrix of B which consists of the $i_1 \dots i_1$-th rows of B. There are three essentially different cases:

(i) $j \notin \{i_1 \dots i_l\}$,

(ii) $j \in \{i_1 \dots i_l\}$ $\quad i \in \{i_1 \dots i_l\}$,

(iii) $j \in \{i_1 \dots i_l\}$ $\quad i \notin \{i_1 \dots i_l\}$.

In the case (i) the matrix $\tilde{B}$ and the submatrix of A which consists of the $i_1 \dots i_l$-th rows of A and is denoted by $\tilde{A}^{i_1 \dots i_l}$ are identical. Therefore $\det \tilde{B} = \det \tilde{A}^{i_1 \dots i_l}$. Though in the case (ii) the matrix $\tilde{B}$ and the matrix $\tilde{A}^{i_1 \dots i_l}$ are not identical, their determinants are, however, equal since the determinant of a matrix does not change if we subtract from one of its rows an other row multiplied by a constant. Thus it is easy to see that $\det \tilde{B} = \det \tilde{A}^{i_1 \dots i_l}$.

Finally, in the case (iii) we assume that $j = i_k$ $(1 < k \leqq l)$. Let us consider the matrices $\tilde{A}^{i_1 \dots i_l}$ and $\tilde{A}^{i_1 \dots i_{k-1}, i, i_{k+1} \dots i_l}$. Using the property of the determinant one can see that

$$\det \tilde{B} = \det \tilde{A}^{i_1 \dots i_l} - c \det \tilde{A}^{i_1 \dots i_{k-1} i i_{k+1} \dots i_l}. \tag{3}$$

Since the matrix $\tilde{A}^{i_1 \dots i_{k-1} i i_{k+1} \dots i_l}$ is not only one of the $l \times l$-form submatrices of A, but of B as well, using the well-known property of g.c.d. by which for all numbers $a, b, c \in \mathbf{Z}$ g.c.d. $(a, b) =$ g.c.d. $(a, b+ca)$ we get the following:

$$\begin{aligned} &\text{g.c.d.}\,(\det \tilde{A}^{i_1 \dots i_{k-1} i i_{k+1} \dots i_l}, \det \tilde{B}) = \\ &= \text{g.c.d.}\,(\det \tilde{A}^{i_1 \dots i_{k-1} i i_{k+1} \dots i_l}, \det \tilde{A}^{i_1 \dots i_l}). \end{aligned} \tag{4}$$

Since for all matrices $\tilde{B}$ belonging to the group (iii) we can find an only matrix $\tilde{A}^{i_1 \dots i_{k-1} i i_{k+1}, \dots, i_l}$ belonging to the group (i) so that the equation (4) holds and keeping in mind the results of the cases (i) and (ii) and that for all numbers $a_1, \dots, a_n \in \mathbf{Z}$ the equation

$$\text{g.c.d.}\,(a_1, \dots, a_n) = \text{g.c.d.}\,(a_1, \dots, \text{g.c.d.}\,(a_i, a_{i+1}), \dots, a_n) \tag{5}$$

holds, we get

$$\text{g.c.d.}\,(\det \tilde{B}, \dots, \det \tilde{B}^{\binom{n}{l}}) = \text{g.c.d.}\,(\det \tilde{A}, \dots, \tilde{A}^{\binom{n}{l}}). \tag{6}$$

(We can pair the determinants in both of the brackets so that finally we have to define the g.c.d. of the same k elements.) This completes the proof of the lemma.

LEMMA 2. *Let $A \in Z^{n \times l}$ be a matrix having integer elements $((\mathbf{a}_{i_j}))$. Then by the help of the row transformation we could see in Lemma 1 the matrix A can be written in the following form:*

$$A' = \begin{pmatrix} a'_{11} & \dots & a_{l_1} \\ 0 & \ddots & \vdots \\ & \ddots & a'_{ll} \\ & & 0 \\ 0 & \dots & 0 \end{pmatrix} \in \mathbf{Z}^{n \times l}. \tag{7}$$

PROOF. Apply the algorithm of Euklid to the n-th and $(n-1)$-th elements of the first column. (If $|a_{1n}| \geqq |a_{1,n-1}|$, then $a_{1,n} = ra_{1,n-1} + q$ where $|q| < |a_{1,n-1}|$, therefore $q = a_{1,n} - ra_{1,n-1}$ and r is determined by the integers $a_{1,n}$ and $a_{1,n-1}$, namely from the n-th row of A we subtract the $(n-1)$-th row of A multiplied by r we say that we have performed the first step of the algorithm of Euklid.) It is easy to see that performing this algorithm we get a matrix like this: the first $(n-2)$ rows of it and the first $(n-2)$ rows of A are identical and the last or the last but one element of the first column of it is equal to zero. We can assume that this element is the last one. (We can change the last two rows if we need to.) Apply this algorithm to the elements $(n-1)$-th and $(n-2)$-th of the first column of the new matrix. It is easy to see that finally we get the matrix A in the following form:

$$A^* = \begin{pmatrix} a_{11}^* & & \dots a_{l1}^* \\ 0 & a_{22}^* & \dots a_{l2}^* \\ \vdots & & \\ 0 & a_{2n}^* & \dots a_{ln}^* \end{pmatrix} \in \mathbf{Z}^{n\times l}. \tag{8}$$

Extending the procedure applying to the n-th and $(n-1)$-th rows and further we get that the form of A is the same as we predicated in (7) which was to be proved.

The proof of the first theorem

It is clear from the definition of the primitive system that a subset of a primitive system is a primitive one as well. The basises of Γ are primitive systems with n elements in Γ, thus if the $\{\mathbf{p}_1, \dots, \mathbf{p}_l\}$ system can be supplemented for a basis of the lattice then it is a primitive system by all means.

Suppose that $\{\mathbf{p}_1, \dots, \mathbf{p}_l\}$ is a primitive system in Γ. Then by Theorem 2 for the $\tilde{A}, \dots, \overset{\tilde{}\binom{n}{l}}{A}$ $l\times l$-form submatrices of coefficient matrix A of the vector system $\{\mathbf{p}_1, \dots, \mathbf{p}_l\}$ the equation (1) holds. By the Lemma 2 the matrix A can be transformed into an "upper triangle matrix" denoted by A' and the g.c.d. of the determinants of the submatrices of the matrix we have got and which has the form (7) is equal by 1 because of the Lemma 1. Since in the matrix talking (7) form only the determinant of the $(l\times l)$-form matrix which is determined by the first l rows is not equal to zero, in this case the determinant of this matrix is equal to 1. The columns of matrix A' are such vectors of the lattice Γ that can be supplemented for a basis of the lattice with the basis vectors $f_{l+1}, \dots, f_n$. (The determinant of the matrix $[A', f_{l+1}, \dots, f_n] \in \mathbf{Z}^{n\times n}$ is equal to 1.)

Let B' be the matrix of the coefficients of this system, namely

$$B' = \begin{pmatrix} a'_{11} & \dots & a'_{l_1} & & \\ & \ddots & \vdots & & 0 \\ & & a'_{ll_1} & & \\ & & & \ddots & \\ 0 & & & & 1 \end{pmatrix} \in \mathbf{Z}^{n\times n}. \tag{9}$$

Let us apply to the matrix B' the inverse of the row-transformation we could see in the Lemma 2. We get such a matrix B that its determinant by the Lemma 1 is equal to 1 and its form is the following:

$$B = \begin{pmatrix} p_1 & \dots p_l & p_{l+1} & \dots & p_n \\ a_{11} & \dots a_{l1} & a_{l+1,1} & & a_{n,1} \\ \vdots & & & & \\ a_{1l} & \dots a_{ll} & a_{ln,l}^{l+1,l} & & a_{n,l} \\ \vdots & & & & \\ a_{1n} & \dots a_{ln} & a_{ln,n}^{l+1} & & a_{n,n} \end{pmatrix} \in \mathbf{Z}^{n\times n}. \tag{10}$$

Thus, if $\{\mathbf{p}_1, \ldots, \mathbf{p}_l\}$ is a primitive system then it can be supplemented for a basis of the lattice Γ with the adequate vectors $\mathbf{p}_{l+1}, \ldots, \mathbf{p}_n$ (writing it in the basis $\{f_1, \ldots, f_n\}$). This completes the proof of the Theorem 1.

The speed of the algorithm

Let us regard a performing of a calculating operator as an elementary step. Denote by $c_1(>0)$ the absolute value of the element of A of which absolute value is not less than the absolute values of the other elements of A. Then the algorithm of Euklid performing with two arbitrary elements of A requires not more than c_1 divisions, thus the first step of the reduction in the Lemma 2 requires not more than $c_1 \cdot l$ multiplications and the same quantity of subtractions. So we have to perform not more than $2c_1 \cdot l(n-1)$ operations for making it be in that form we need. For this reason the number of the operations that we have to perform transforming A into A' is not more than

$$2c_1[l(n-1)+(l-1)(n-2)+\ldots+1\cdot(n-1)]. \tag{11}$$

In a similar manner it is easy to see that the number of those operations that we have to perform transforming B' into B is not more

$$2c_1[(n-1)(n-l-1)+(n-l+1)(n-l)+\ldots+n(n-1)]. \tag{12}$$

Since we can make an upper estimating for both of the (11) and (12) forms by the following:

$$\begin{aligned} 2c_1[n(n-1)+\ldots+2\cdot 1] &= 2c_1\Big[\sum_{k=1}^{n} k^2 - \sum_{k=1}^{n} k\Big] = \\ = 2c_1\left[\frac{n(n+1)(2n+1)}{6} - \frac{n(n+1)}{2}\right] &= 2c_1\frac{n(n+1)(n-1)}{3}, \end{aligned} \tag{13}$$

therefore the number of the steps we require is not more than

$$c(n-1)n(n+1) \qquad (\text{viz.} = O(n^3)), \tag{14}$$

where c is a constant depending only on the coefficients of the primitive system. Thus the algorithm is a polynomial function of the number of the dimension.

REFERENCES

[1] Waerden, B. L. van der, Die Reduktionstheorie der positiven quadratischen Formen, *Acta Math.* **96** (1956), 265—309. *MR* **18**—562.

[2] G. Horváth, Ákos, The algebraic characterization of the primitive systems, *Studia Sci. Math. Hungar.* **24** (1989), 331—337.

(Received February 9, 1987)

BUDAPESTI MŰSZAKI EGYETEM
GEOMETRIAI TANSZÉK
SZTOCZEK U. 4
H—1521 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 331—337

ALGEBRAIC CHARACTERIZATION OF PRIMITIVE SYSTEMS

ÁKOS G. HORVÁTH

Introduction

One of the most important concept of basis reduction problems of n-dimensional lattice geometry is the definition of primitive system. The geometrical characterization of these systems is given by van der Waerden in his fundamental work [1]. He proved that a vector system in the lattice is primitive system if and only if it can be complemented for a basis of the lattice. The purpose of this paper is to give an algebraic characterization of these systems. By the help of this characterization the van der Waerden's theory can be proved and it can be applied well to the solution of a lot of reduction problems.

Definitions, notations

E^n: Euclidean n-space
$\mathbf{Z}$: the set of integers
$\mathbf{Z}^{n\times k}$: the set of integer elemented matrices with n rows and k columns
$\mathbf{Q}$: the set of rational numbers
$\mathbf{R}$: the set of real numbers
a/b: a divides b
g.c.d. $(x_1, \ldots, x_n)$ is the greatest common divisor of the integer numbers $x_1, \ldots, x_n$
n-lattice: Let $\mathbf{r}_1, \ldots, \mathbf{r}_n$ be linearly independent vectors in E^n. Let $[\mathbf{r}_1, \ldots, \mathbf{r}_n] = \{\mathbf{x}\in E^n;\ \mathbf{x}=\sum_{i=1}^{n} \mathbf{x}_i\mathbf{r}_i\ \forall i\ 1\leqq i\leqq n,\ x_i\in\mathbf{Z}\}$. $[\mathbf{r}_1, \ldots, \mathbf{r}_n]$ is called n-lattice with basis $\{\mathbf{r}_i\}$, $i=1, \ldots, n$
primitive system: Let Γ be an n-lattice, and let $\mathbf{a}_1, \ldots, \mathbf{a}_k$ $(1\leqq k\leqq n)$ be linearly independent vectors of Γ. We say that $\{\mathbf{a}_i\}$, $i=1, \ldots, k$ is a primitive system in Γ if it is a basis of the sublattice lying in their subspace. Denote by $L[\mathbf{a}_1, \ldots, \mathbf{a}_k]$ the subspace which is spanned by $\{\mathbf{a}_i\}$. Obviously, $\{\mathbf{a}_i\}$ is primitive system in Γ iff the following condition holds:

$$[\mathbf{a}_1, \ldots, \mathbf{a}_k] = \Gamma\cap L[\mathbf{a}_1, \ldots, \mathbf{a}_k] \tag{1}$$

primitive vector: primitive system consisting of one element.
The following statement can be proved easily:
all subsets of a primitive system are also primitive ones.

1980 *Mathematics Subject Classification* (1985 Revision). Primary 10E25; Secondary 10E05.
Key words and phrases. Lattice, primitive system.

The theorem

Let $A \in \mathbf{Z}^{n\times k}$ be arbitrary matrix. Denote by $\tilde{A}, \ldots, \overset{\binom{n}{k}}{\tilde{A}}$ the k-dimensional determinants of matrix A.

THEOREM. *Let $\Gamma = [\mathbf{f}_1, \ldots, \mathbf{f}_n]$ be an arbitrary n-lattice and p_i, $i=1, \ldots, k$ are linearly independent vectors in Γ. If we denote by A the matrix of the coefficients of $\{\mathbf{p}_i\}$, $i=1, \ldots, k$ then $A \in \mathbf{Z}^{n\times k}$ and the following statements are equivalent:*

(2)

(i) $\{\mathbf{p}_1, \ldots, \mathbf{p}_k\}$ *is a primitive system*

(ii) g.c.d. $(\det \tilde{A}, \ldots, \det \overset{\binom{n}{k}}{\tilde{A}}) = 1$ *(namely the greatest common divisor of the k-dimensional determinants of A is equal to one.)*

Lemmas

LEMMA 1. *Let $\mathbf{p}_1, \ldots, \mathbf{p}_k$ be vectors in $\Gamma = [\mathbf{f}_1, \ldots, \mathbf{f}_n]$. If $\mathbf{q} = \sum_{i=1}^{k} \alpha_i \mathbf{p}_i$ is also lattice vector then $\alpha_1, \ldots, \alpha_k \in \mathbf{Q}$.*

PROOF. Since the vectors $\{\mathbf{p}_i\}$ are lattice vectors they can be written by means of vectors $\mathbf{f}_1, \ldots, \mathbf{f}_n$ as follows:

$$\begin{aligned} \mathbf{p}_1 &= a_{11}\mathbf{f}_1 + \ldots + a_{1n}\mathbf{f}_n \\ &\vdots \\ \mathbf{p}_k &= a_{k1}\mathbf{f}_1 + \ldots + a_{kn}\mathbf{f}_n \end{aligned} \tag{3}$$

where $a_{ij} \in \mathbf{Z}$. This means that $\mathbf{q} = \alpha_1\mathbf{p}_1 + \ldots + \alpha_k\mathbf{p}_k = (\alpha_1 a_{11} + \ldots + \alpha_k a_{k1}) f_1 + \ldots \ldots + (\alpha_1 a_{1n} + \ldots + \alpha_k a_{kn}) f_n$. Since $\mathbf{q}$ is in Γ there are integer numbers $\mathbf{d}_i$, $i=1, \ldots, n$ for which $\{\alpha_1, \ldots, \alpha_k\}$ is solution of the following system of equations:

$$\begin{aligned} a_{11}\alpha_1 + \ldots + a_{k1}\alpha_k &= d_1 \\ \\ a_{1n}\alpha_1 + \ldots + a_{kn}\alpha_k &= d_n. \end{aligned} \tag{4}$$

Let us suppose that we know the numbers $\{d_i\}$. Then by solving the system of equations (4) by the help of Gauß's elimination it is easy to see that the numbers $\{\alpha_i\}$ are rationals.

LEMMA 2. *Let $\mathbf{p}_1, \ldots, \mathbf{p}_k$ be vectors in $\Gamma = [\mathbf{f}_1, \ldots, \mathbf{f}_n]$. The system $\{\mathbf{p}_i\}$ is a primitive one in Γ if and only if arbitrary numbers $d_1 \ldots d_n \in \mathbf{Z}$ on which system of equations* (4) *can be solved the solution is integer, namely $\alpha_1, \ldots, \alpha_k \in \mathbf{Z}$.*

PROOF. Assume that the system $\{\mathbf{p}_1, \ldots, \mathbf{p}_k\}$ is a primitive one. From the definition of the primitive system follows directly that $\{\mathbf{p}_i\}$ is basis of this k-sublattice of Γ which is lying in k-dimensional linear subspace spanned by $\{\mathbf{p}_i\}$. Therefore if the lattice vector $\mathbf{q} = d_1\mathbf{f}_1 + \ldots + d_n\mathbf{f}_n$ can be written with real linear combination of vectors $\{\mathbf{p}_i\}$ then it can be written with integer coefficient, too. This means that the only solution $\alpha_1, \dots, \alpha_k$ of the system of equation (4) is integer, namely $\alpha_1, \dots, \alpha_k \in \mathbf{Z}$. If the vector $\mathbf{q}$ cannot be written as linear combination of vectors $\{\mathbf{p}_i\}$ then (4) has no solution.

Conversely, assume that all such integer numbers $d_1, \dots, d_n$ on which (4) can be solved, the solution is integer. Then each lattice-vector $\mathbf{q}$ which can be written as real linear combination of vectors $\{\mathbf{p}_i\}$ is in the sublattice $[\mathbf{p}_1, \dots, \mathbf{p}_k]$. Therefore $\mathbf{p}_1, \dots, \mathbf{p}_k$ is basis in lattice $\Gamma \cap L[\mathbf{p}_1, \dots, \mathbf{p}_k]$, namely $\mathbf{p}_1, \dots, \mathbf{p}_k$ is primitive system in Γ.

Let A be the coefficient matrix of system of equations (4), and let $A_i \in \mathbf{Z}^{n \times (k-1)}$ $(1 \leqq i \leqq k)$ be that matrix which is derived from A by omitting the i-th column.

LEMMA 3. *Let $1 \leqq i \leqq k$ be arbitrary integer. Assume that the greatest common divisor of the $(k-1)$-dimensional determinants of A_i is greater than one. Then the greatest common divisor of the $k \times k$ determinants of A is also greater than one.*

PROOF. Without loss of generality we can assume that $i=k$. Select that submatrix of A which is defined by the $i_1, \dots, i_k$-th rows of A. The determinant of this submatrix can be written in the following form:

$$\det\begin{pmatrix} a_{1i_1} & \dots & a_{ki_1} \\ \vdots & & \\ a_{1i_k} & \dots & a_{ki_k} \end{pmatrix} = a_{ki_1} \det(A_{ki_1}) + \dots + a_{ki_k} \det(A_{ki_k}) \tag{5}$$

where $\det(A_{ki_j})$ is the relevant subdeterminant of the selected submatrix. According to the condition $\det(A_{ki_j})$ can be divided with an integer which is greater than one therefore the left-hand side of (5) can also be divided with this integer. Regarding the fact that arbitrary k-dimensional submatrices of A have this property, it is clear that the statement of Lemma 3 is true.

LEMMA 4. *Assume that the greatest common divisor of the $(k-1)$-dimensional determinants of A_i is equal to one where $1 \leqq i \leqq k$ is arbitrary integer. Let us suppose further that there are such $d_1, \dots, d_n \in \mathbf{Z}$ numbers that the system of equations* (4) *can be solved by numbers $\alpha_1, \dots, \alpha_k \in \mathbf{R}$. Then α_i can be written in form $\dfrac{x_i}{y}$ where $y, x_i \in \mathbf{Z}$ $i=1, \dots, k$ and* g.c.d.$(x_i | y) = 1$.

PROOF. Using the Lemma 1. We get that α_i can be written in the form $\dfrac{x_i}{y_i}$ where g.c.d.$(x_i | y_i) = 1$ $(x_i, y_i \in \mathbf{Z}, \forall i = 1, \dots, k)$. Substituting these numbers into (4) we get the following system of equations:

$$\begin{aligned} a_{11} x_1 y_2 \dots y_k + \dots + a_{k1} y_1 \dots y_{k-1} x_k &= d_1 y_1 \dots y_k \\ &\vdots \\ a_{1n} x_1 y_2 \dots y_k + \dots + a_{kn} y_1 \dots y_{k-1} x_k &= d_n y_1 \dots y_k \end{aligned} \tag{6}$$

Since the product $y_i \dots y_{k-1}$ divide the right-hand side of an the last member of the left-hand side of the all equations of (6) therefore the following conditions hold:

$$\begin{aligned} y_1 \dots y_{k-1} &| a_{11} x_1 y_2 \dots y_k + \dots + a_{k-1,1} y_1 \dots y_{k-1} y_k =: c_1 \\ &\vdots \\ y_1 \dots y_{k-1} &| a_{1k-1} x_1 y_2 \dots y_k + \dots + a_{k-1,k-1} y_1 \dots y_{k-2} x_{k-2} y_k = c_{k-1}. \end{aligned} \tag{7}$$

Denote by B_{ij} the $(k-2)$-dimensional subdeterminant of $\bar{A}$, where

$$\bar{A} := \begin{pmatrix} a_{11} \dots a_{k-1,1} \\ a_{1,k-1} \dots a_{k-1,k-1} \end{pmatrix} \in \mathbf{Z}^{(k-1)\times(k-1)},$$

which belongs to element a_{ij} of $\bar{A}$. By the help of (7) we get that

(8) $$y_1 \cdot \ldots \cdot y_{k-1} | c_1 B_{11} + \ldots + C_{k-1} B_{1,k-1}.$$

Using the expansion and skew-expansion theorems we get the following relation:

(9) $$y_1 \cdot \ldots \cdot y_{k-1} | (\det \bar{A})\, x_1\, y_2 \ldots y_k,$$

namely

(10) $$y_1 | (\det \bar{A})\, x_1\, y_k$$

holds. Set out from another rows of (6) by similar arguments we get:

(11) $$y_1 | (\det \bar{\bar{A}})\, x_1\, y_k$$

where $\bar{\bar{A}}$ is defined in a similar manner than $\bar{A}$. Respecting the fact that it can be said for arbitrary $(k-1)$ rows of (6) we get the following conditions:

(12) $$y_1 | (\det \bar{A})\, x_1 y_k \quad y_1 | (\det \bar{\bar{A}})\, x_1 y_k \ldots y_1 | (\det \overset{\bar{\vdots}\binom{n}{k-1}}{A})\, x_1 y_k$$

where $\det \bar{A}, \ldots, \det \overset{\bar{\vdots}\binom{n}{k-1}}{A}$ are all of the $(k-1)$-dimensional subdeterminants of A_k. Using the assumption of Lemma 4 g.c.d.$((\det \bar{A}), \ldots, (\det \overset{\bar{\vdots}\binom{n}{k-1}}{A})) = 1$, therefore $y_1 | x_1 \cdot y_k$. Since g.c.d.$(x_1 | y_1) = 1$ we get that

(13) $$y_1 | y_k.$$

If we study the divisibility of (6) by $y_2 \cdot \ldots \cdot y_k$ we can see that

(14) $$y_k | y_1.$$

We can assume that the signs of integers $y_1, \ldots, y_k$ equal to each other therefore $y_1 = y_k$. By similar arguments it is easy to see that $y_i = y_j$. This completes the proof of the lemma.

The proof of the theorem

The proof is carried out by induction with respect to the number of elements of the system. In the case $k=1$ the statement is obvious. Assume that $n \geqq k \geqq 1$ and the statement is true in the case $1 < k$. First we verify that if the condition of theorem is satisfied by the system $\mathbf{p}_1, \ldots, \mathbf{p}_k$ then it is a primitive one (1).

Secondly, we shall prove if the condition is not satisfied by $\mathbf{p}_1, \ldots, \mathbf{p}_k$ then it is notprimitive system (2).

1. Assume that g.c.d.$(\det \tilde{A}, \ldots, \det \overset{\tilde{\vdots}\binom{n}{k}}{A}) = 1$. Using Lemma 3 we get that the greatest common divisor of all $(k-1)$-dimensional subdeterminants of

$A_i \in \mathbf{Z}^{n\times(k-1)}$ ($1 \leqq i \leqq k$ arbitrary integer) is equal to one. Study again the system of equations (4)! Let us suppose indirectly that $\{\mathbf{p}_1, \ldots, \mathbf{p}_k\}$ is not a primitive system. Using Lemma 1 and Lemma 2, we get that there are such numbers $d_1, \ldots, d_n \in \mathbf{Z}$ and $\alpha_1, \ldots, \alpha_k \in \mathbf{Q}$, which satisfy the system of equations (4) and at least one of the $\{\alpha_i\}$ is not integer. Since the conditions of Lemma 4 hold the system of equations (4) can be written in the following form:

$$\begin{aligned} a_{11}x_1 + \ldots + a_{k1}x_k &= d_1 y \\ &\vdots \\ a_{1n}x_1 + \ldots + a_{kn}x_n &= d_n y. \end{aligned} \tag{15}$$

Let $\tilde{A}$ be that submatrix of A which is determined by the first k rows of A. Denote by C_{ij} the subdeterminant of $\tilde{A}$ which belong to the j-th element of the i-th row. If we multiply the i-th row of (15) by C_{1i} ($1 \leqq i \leqq k$) and we sum up these products for the first k's, then we get:

$$\begin{aligned} (C_{11}a_{11} + \ldots + C_{1k}a_{1k})x_1 + \ldots + (C_{11}a_{k1} + \ldots + C_{1k}a_{kk})x_k &= \\ = y(C_{11}d_1 + \ldots + C_{1k}d_k). \end{aligned} \tag{16}$$

Using the expansion and skew-expansion theorems we find:

$$(\det \tilde{A})x_1 = y(C_{11}d_1 + \ldots + C_{1k}d_k), \tag{17}$$

however, g.c.d.$(x_1 | y) = 1$, therefore $y | \det \tilde{A}$. Setting out from another k rows of A by similar arguments we get

$$y|\det \tilde{A} \ldots y|\det \tilde{A}^{\binom{n}{k}} \tag{18}$$

i.e.

$$y|\text{g.c.d.}\,(\det \tilde{A}, \ldots, \det \tilde{A}^{\binom{n}{k}}) = 1. \tag{19}$$

Since the relation (19) immediately follows that $\alpha_1, \ldots, \alpha_k$ are integers, we get contradiction with the initial assumption, namely $\mathbf{p}_1, \ldots, \mathbf{p}_k$ is primitive system.

2. Let us suppose now that the condition concerning the subdeterminants of A is not satisfied, namely $y := \text{g.c.d.}(\det \tilde{A}, \ldots, \det \tilde{A}^{\binom{n}{k}}) > 1$. We shall prove that the system $\{\mathbf{p}_1, \ldots, \mathbf{p}_k\}$ is not primitive. In the sense of Lemma 2 we ought to find such rational numbers $\alpha_1, \ldots, \alpha_k$ which substituting for the system of equations (4) the values of left-hand side of (4) are integer numbers and at least one of the $\{\alpha_i\}$ is not integer. If system $\{\mathbf{p}_1, \ldots, \mathbf{p}_{k-1}\}$ is not primitive then the system $\mathbf{p}_1, \ldots, \mathbf{p}_k$ also may not be primitive system. We can assume that $\{\mathbf{p}_1, \ldots, \mathbf{p}_{k-1}\}$ is a primitive one. Using the condition of induction the greatest common divisor of all $(k-1)$-dimensional determinants of matrix A_k of this system is equal to one. For this reason there is such a $(k-1)$-dimensional subdeterminant of A_k which is not divisible by y. Without loss of generality we can assume that it is determined by the first $(k-1)$ rows of A_k. Denote by D_k this subdeterminant. Let $B \in \mathbf{Z}^{(k-1)\times k}$ be the submatrix of A which is

determined by these $(k-1)$ rows. Let D_i be that $(k-1)$-dimensional subdeterminant of B which we find by omitting the i-th column of B. Define the numbers α_i as follows:

$$\alpha_i = \frac{D_i}{y}, \quad i = 1, \dots, k. \tag{20}$$

At this time the system (4) can be written in the following form:

$$\begin{aligned} a_{11}\frac{D_1}{y}+\dots+a_{k1}\frac{D_k}{y} &= d_1 \\ a_{1k-1}\frac{D_1}{y}+\dots+a_{kk-1}\frac{D_k}{y} &= d_{k-1} \\ a_{1k}\frac{D_1}{y}+\dots+a_{kk}\frac{D_k}{y} &= d_k \\ &\vdots \\ a_{1n}\frac{D_1}{y}+\dots+a_{kn}\frac{D_k}{y} &= d_n. \end{aligned} \tag{21}$$

If the index "j" is less than "k", then the value of the left-hand side of the j-th row is equal to zero (It is easy to see by the help of the skew-expansion theorem), namely $d_1=\dots=d_{k-1}=0\in\mathbf{Z}$. If the index j is greater than $(k-1)$, then the value of the left-hand side of the j-th row is equal to a k-dimensional determinant divided by y of A, $\left(\text{namely } \dfrac{\det \tilde{A}}{y}\right)$ therefore it is integer number, $d_k, \dots, d_k\in\mathbf{Z}$. We have to prove that at least one of these $\{\alpha_i\}$ is not integer. However, in the sense of the definition of D_k it is obvious that $\alpha_k=\dfrac{D_k}{y}\in\mathbf{Z}$. Therefore if the condition concerning the subdeterminants of A is not satisfied then the system $\{\mathbf{p}_1, \dots, \mathbf{p}_k\}$ is not primitive. Thus the theorem is proved.

Remark

By the help of the van der Waerden's theorem the necessity of the condition concerning the subdeterminants can be simply proved as follows. Assume that the greatest common divisor of the k-dimensional determinants of $A\in\mathbf{Z}^{n\times k}$ is equal to p. (The matrix A is the matrix of the coefficients of the system $\{\mathbf{p}_1, \dots, \mathbf{p}_k\}$.) Complete the matrix A to an integer matrix $M\in\mathbf{Z}^{n\times n}$. By the help of expansion theorem of Laplace the number $\det M$ can be written in the following form:

$$\det M = q_1(\det \tilde{A})+\dots+q_{\binom{n}{k}}\left(\det \tilde{A}^{\binom{n}{k}}\right), \tag{22}$$

where $q_i, i=1, \ldots, \binom{n}{k}$ are integers. It is clear that $p|\det M$. If $\{\mathbf{p}_1, \ldots, \mathbf{p}_k\}$ was primitive system in Γ then using van der Waerden's theorem we get that it can be complemented for a basis of the lattice, namely there is such an integer matrix $M \in \mathbf{Z}^{n \times n}$ which contains the matrix $A \in \mathbf{Z}^{n \times k}$ and its determinant is equal to one. Considering the equation (22) we find that $p|\det M=1$ namely $p=1$.

Similar simple proof of the sufficiency of the condition the author does not know. However, by the help of this characterization can be proved the van der Waerden's theorem. It can be read in [2].

Notice that the trivial statement of Lemma 1 follows from the theory of simple algebraic extension of the rational numbers, since one can obtain the one-to-one correspondence between the lattice Γ and the set of algebraic integers of order n, if the base $\{e_1, \ldots, e_n\}$ of the lattice Γ corresponds to the base $\{1, v, \ldots, v^{n-1}\}$ of the simple algebraic extension of order n of the rational numbers. Using the fact that $n+1$ of the set of algebraic numbers of order n must be connected over the rational numbers ([3]), so $n+1$ of the lattice vectors is connected over $\mathbf{Q}$, as well. The proof given in the paper for Lemma 1 is, however, independent from the theory of simple algebraic extension of rational numbers.

REFERENCES

[1] Waerden, B. L. van der, Die Reduktionstheorie der positiven quadratischen Formen, *Acta Math.* **96** (1956), 265—309. *MR* **18**—562.
[2] G. Horváth, Á., On a polynomial algorithm for selecting a lattice-basis containing a given primitive system, *Studia Sci. Math. Hungar.* **24** (1989), 325—329.
[3] Landau, E., *Einführung in die elementare und analytische Theorie der algebraischen Zahlen und der Ideale*, Leipzig und Berlin, 1918.

(Received February 9, 1987)

BUDAPESTI MŰSZAKI EGYETEM
GEOMETRIA TANSZÉK
SZTOCZEK U. 4
H—1521 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 339—344

A NOTE ON THE STOCHASTIC GEYSER PROBLEM

KARL GRILL

Abstract

Let $X_1, X_2, \ldots$ be i.i.d.r.v. and $S_n = X_1 + X_2 + \ldots + X_n$ their cumulative sums. Suppose that we observe that quantities $\hat{S}_n = S_n + R_n$ where R_n is a small random perturbation. Bártfai (1966) proved that if the moment generating function (m.g.f.) of X_1 exists, and if $R_n = o(\log n)$ a.s., then the distribution function of X_1 can be determined with probability one from the infinite sequence of observations.

His proof, however, (and also the ones given later, cf., e.g., Erdős—Rényi 1970) does not give a clue as to how one could estimate the distribution function (or a functional thereof) from a finite set of observations.

In the present paper, we shall give a sequence of estimators for the m.g.f. of X_1 which will then be used to derive estimators for the k-th moments that will turn out to be valid even in the case when the m.g.f. itself does not exist.

1. Introduction

Rényi (1962) posed the following problem: Suppose that Robinson Crusoe has a geyser on his island, and that the time intervals between the eruptions of that geyser are i.i.d. bounded r.v.. Having recorded the number of eruptions per day for a long time, Robinson decides that it should be possible to estimate the distribution of the time between two eruptions.

In mathematical terms, this *stochastic geyser problem* restates as follows: Let $X_1, X_2, \ldots$ be i.i.d.r.v. and $S_n = X_1 + \ldots + X_n (S_0 = 0)$. Suppose further that the X_n are positive and bounded. Can we then determine the distribution function of X_1 from the sequence $[S_n]$?

Bártfai (1966) proved the somewhat more general

THEOREM A. *Let $X_1, X_2, \ldots$ be i.i.d.r.v. (not necessarily bounded) and suppose that the m.g.f. $M(t) = \mathsf{E} \exp(X_1 t)$ exists for t in some interval around the origin. Suppose further that R_n is a sequence of random variables with $R_n = o(\log n)$ a.s.. Then from the infinite sequence $\hat{S}_n = S_n + R_n$ one can determine the distribution function of X_1 with probability one.*

Bártfai's proof, as well as another one given by Erdős and Rényi (1970) is based on Petrov's (1969) large deviation theorem and gives no direct clue as to how one can construct estimators for the distribution function or a functional thereof — e.g., the moments — based on a finite number of observations.

1980 *Mathematics Subject Classification* (1985 Revision). Primary 62G05; Secondary 62G30, 60G50.

Key words and phrases. Estimation, moments, moment generating function.

In the sequel, we shall give a sequence of estimators for the m.g.f. $M(t)$; unfortunately, the distribution function itself cannot be handled that easily. It will turn out, however, that our estimators are quite good (in a certain sense) and that they are also suitable for estimating the moments of X_1. In fact, one can consistently estimate the moments in this way even if the m.g.f. does not exist.

2. Estimating the m.g.f.

The m.g.f. being defined as an expectation, it would seem the most natural course to estimate it by taking time averages. In our setting, however, this direct course is not feasible because of the presence of the perturbations R_n. Yet, observing that the m.g.f. of a sum of k consecutive X_n's is the k-th power of the m.g.f. of X_1, one might try to estimate the former by taking time averages, then take the k-th root to obtain the latter. This programme gives rise to two sensible types of estimators:

$$\hat{M}_n(t) = \left(\frac{k}{n}\sum_{j=1}^{[n/k]-1} \exp\left((\hat{S}_{(j+1)k}-\hat{S}_{jk})\,t\right)\right)^{1/k}$$

$$\overline{M}_n(t) = \left(\frac{1}{n}\sum_{j=0}^{n/k} \exp\left((\hat{S}_{j+k}-\hat{S}_{j})\,t\right)\right)^{1/k}.$$

These estimators, of course, implicitly depend on the choice of k as a function of n. In making this choice, one has to achieve two conflicting goals: On the one hand, k has to be large enough that by taking the root the perturbations R_n are sufficiently reduced, on the other hand, it has to be small enough that the effect of the intrinsic randomness of the sequence (X_n) is concealed by the averaging process. It will turn out that $k_n \approx C \log n$ is a suitable choice.

Now, although the estimator $\overline{M}_n$ seems to have a somewhat smaller variation, we shall only study the performance of $\hat{M}_n$, since this one is easier to handle mathematically, and since in the scope of our investigation there is no difference between these two kinds of estimators.

Let us now state our main result:

THEOREM 1. *Let $X_1, X_2, \ldots$ be i.i.d.r.v. and let I_0 be the largest open interval where $M(t)$ exists. Furthermore, assume that there is a nondecreasing sequence C_n such that $R_n \leqq C_n$ for n big enough. Then it holds:*

$$\hat{M}_n(t) = M_n(t)\left(1+O\left(\frac{C_n}{\log n}\right)\right) \tag{1}$$

uniformly on compact subintervals of I_0.

REMARK 1. As will be clear from the proof of our theorem, one can get the same rate of convergence also in the case where C_n tends to 0 if one restricts ones attention to small subintervals of I_0.

REMARK 2. The rate obtained in (1) is best possible: this is an almost trivial consequence of the Komlós—Major—Tusnády (1975) approximation scheme.

PROOF. Without loss of generality, we may restrict ourselves to positive t. Furthermore, it is clear that

$$\hat{M}_n(t) = \tilde{M}_n(t)\left[1+O\left(\frac{R_n}{k_n}\right)\right]$$

with

$$\tilde{M}_n(t) = \left(\frac{k}{n}\sum_{j=1}^{[n/k]-1} \exp\left((S_{(j+1)k}-S_{jk})\,t\right)\right)^{1/k}.$$

So we shall investigate the behaviour of $\tilde{M}_n(t)$ as $n\to\infty$. Let us first consider a single fixed $t\in I_0$. We can clearly find an $\varepsilon>0$ such that $(1+\varepsilon)t\in I_0$. Letting

$$Y_j = \exp\left(S_{(j+1)k}-S_{jk}\right)t)$$

this implies that

$$\mathsf{E}Y_j^{1+\varepsilon}$$

exists. Since it is also clear that

$$\mathsf{E}Y_j = M^k(t).$$

Now, if $2t\in I_0$, we can use Kolmogorov's inequality to obtain

$$\mathsf{P}\left(\max_{j\le n}|\tilde{M}_n^k(t)-M^k(t)| \le \frac{1}{2}M^k(t)\right) \le \frac{4k\,M^k(2t)}{nM^{2k}(t)}.$$

If t is larger, then we proceed as follows: Let

$$\tilde{Y}_j = \begin{cases} Y_j & \text{if } \;|Y_j| \le \dfrac{n}{k}\,\mathsf{E}Y_j \\ 0 & \text{otherwise.} \end{cases}$$

It is readily obtained that

$$\mathsf{P}(\exists j \le n\colon\ \tilde{Y}_j \ne Y_j) \le \frac{k^\varepsilon M^k((1+\varepsilon)\,t)}{n^\varepsilon M^{(1+\varepsilon)k}(t)},$$

$$|\mathsf{E}_j\,\tilde{Y}-\mathsf{E}Y_j| \le \left(\frac{k}{n}\right)^\varepsilon M^k((1+\varepsilon)\,t),$$

and

$$\mathsf{E}\tilde{Y}_j^2 \le \left(\frac{n}{k}\right)^{1-\varepsilon} M^k((1+\varepsilon)\,t).$$

Using again the Kolmogorov inequality on $\tilde{Y}_j$, we finally obtain

$$\mathsf{P}\left(\max_{j\le n}|\tilde{M}_n^k(t)-M^k(t)| \le \frac{1}{2}M^k(t)\right) \le \frac{5k^\varepsilon M^k((1+\varepsilon)t)}{n^\varepsilon M^{(1+\varepsilon)k}(t)\left(1-\dfrac{2k^\varepsilon M^k((1+\varepsilon)t)}{n^\varepsilon M^{(1+\varepsilon)k}(t)}\right)}.$$

Choosing now $k=\delta\log n$ with δ sufficiently small, we obtain that

$$\mathsf{P}\left(\max_{j\leq n}|\tilde{M}_n^k(t)-M^k(t)|\leq\frac{1}{2}M^k(t)\right)\leq n^{-\eta}.$$

Now in order to get a uniform estimate, let $t_j^{(n)}=j/k$ and fix a compact subinterval $I\subset I_0$. From the above calculations it is clear that

$$\mathsf{P}\left(\exists t_l^{(n)}\in I\colon \max_{j\leq n}|\tilde{M}_k^n(t_l^{(n)})-M^k(t_l^{(n)})|\leq\frac{1}{2}M^k(t_l^{(n)})\right)\leq C\,kn^{-\eta}.$$

Let now $n_r=[2^{r/2}]$. It follows from the Borel—Cantelli lemma that with probability one for all r big enough and for all $t_l^{(n_r)}\in I$

$$\max_{j\leq n_r}|\tilde{M}_j^k(t_l^{(n_r)})-M^k(t_l^{(n_r)})|\leq\frac{1}{2}M^k(t_l^{(n_r)}).$$

For arbitrary n we can find an r such that $n_{r-1}<n<n_r$. The above inequality implies that for $t_l^{(n_r)}\in I$

$$\tilde{M}_n(t_l^{(n_r)})=M^k(t_l^{(n_r)})\left[1+O\left(\frac{1}{\log n}\right)\right].$$

For arbitrary $t\in I$ the same relation follows from a simple application of Hölder's inequality, which completes the proof of Theorem 1.

3. Estimating the moments

Now that we have obtained an estimate for $M(t)$, it seems plausible that by some kind of approximate differentiation (i.e., replacing the differential operator by an appropriate finite difference), we could get an estimator for the moments of X_1. This is really the case, and, as it has already been stated, the use of that kind of estimator is not limited to the case when the m.g.f. exists.

For the time being, however, let us assume that the m.g.f. exists. In this case, we have

THEOREM 2. *Let* $X_1, X_2, \ldots$ *satisfy the conditions of Theorem 1 with* $C_n=o(\log n)$. *Furthermore, let* $k=\varepsilon\log n$ *and* $\theta=\left(\frac{C_n}{\log n}\right)^{\varepsilon}$. *Then it holds:*

$$\hat{m}_r=\theta^{-r}\sum_{j=1}^{r}(-1)^{r-j}\binom{r}{j}\hat{M}_n(j\theta) \tag{2}$$

is a strongly consistent estimator of the r-th moment m_r *of* X_1 *if* ε *is sufficiently small.*

PROOF. From the proof of Theorem 1 we know that we can achieve $|\hat{M}_n-M|\leq K\dfrac{C_n}{\log n}$ for n big enough. Thus, if we replace $\hat{M}_n$ by M_n in (2), this will change the result by at most $O\left(\dfrac{C_n}{\theta^n\log n}\right)$. On the other hand, the so modified sum is nothing but the r-fold difference operator, applied to $M(t)$ at $t=0$ and clearly differs from $M^{(r)}(0)=m_r$ by at most $O(\theta)$. So, putting things together, Theorem 2 readily follows.

REMARK. Going to a little more detail, one can also obtain convergence rates in Theorem 2. In fact, by replacing the sum in (2) by a higher-order approximation of the r-th derivative, we immediately find an error term of order $\left(\frac{C_n}{\log n}\right)^{1-\varepsilon}$ for arbitrary $\varepsilon>0$. For $r\geqq 3$ this is almost the best one we can achieve; for $r=2$ our methods can be used to obtain an error term of order $\left(\frac{C_n}{\sqrt{n}}\right)^{1/2+\varepsilon}$. This means that the variance can be estimated in this way whenever $C_n=o(\sqrt{n})$.

Finally, let us assume that we only know that m_r exists. In this case our estimating scheme can also be applied, although with a slight modification. Namely, let us define

$$K_n = \left\{j \leqq \frac{n}{k} : \forall i : jk+1 \leqq i \leqq (j+1)k : \hat{S}_i - \hat{S}_{i-1} \leqq 2C_n\right\}$$

and

$$\hat{\hat{M}}_n(t) = \left(\frac{1}{\#(K_n)} \sum_{j\in K_n} \exp\left((\hat{S}_{(j+1)k} - \hat{S}_{jk})\, t\right)\right)^{1/k}.$$

In plain speech, this means that we discard blocks that contain too large X_i's. With this adjustment, it holds:

THEOREM 3. *If m_r exists, $C_n=o(\log^{1/r} n)$, and if we choose*

$$k = \varepsilon \log n,$$

and

$$\theta = k^{-1/r},$$

then the estimator $\hat{m}_r$ given in (2) is strongly consistent if in its definition $\hat{M}$ is replaced by $\hat{\hat{M}}$ and if ε is sufficiently small.

For a proof, observe that the modification in our estimator $\hat{M}$ amounts to truncating the distribution function of X_j at some random point between C_n and $3C_n$. An easy calculation yields that for the values of t in question the effect of the precise location of this truncation point on the m.g.f. of the truncated variable is of smaller order than t^r; using this, our theorem follows by the same kind of simple estimates as we used in Section 2. Since the detailed calculations do not contain any new ideas, they will be omitted.

Of course, in assuming the existence of higher moments, one can obtain convergence rates, too, and also lessen the restrictions imposed on C_n, but this gets a bit tedious, so we shall not go into it.

REFERENCES

[1] BÁRTFAI, P., Die Bestimmung der zu einem wiederkehrenden Prozess gehörenden Verteilungsfunktion aus den mit Fehlern behafteten Daten einer einzigen Realisation, *Studia Sci. Math. Hungar.* **1** (1966), 161—168. *MR* **35** # 6218

[2] ERDŐS, P. and RÉNYI, A., On a new law of large numbers, *J. Analyse Math.* **23** (1970), 102—111. *MR* **42** # 6907.

[3] KOMLÓS, J., MAJOR, P. and TUSNÁDY, G., An approximation of partial sums of independent RV's and the sample DF.I, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **32** (1975), 111—131. *MR* **51** # 11605b.

[4] RÉNYI, A., A sztochasztikus gejzir (abstract of lecture), *Magyar Tud. Akad. Mat. Kut. Int. Közl.* **7** (1962), 643.

[5] PETROV, V. V., On the probabilities of large deviations of sums of independent random variables, *Theory Probab. Appl.* **10** (1969), 287—298.

(Received April 16, 1987)

INSTITUT FÜR STATISTIK UND
WAHRSCHEINLICHKEITSTHEORIE
TECHNISCHE UNIVERSITÄT WIEN
WIEDNER HAUPTSTRASSE 8—10
A—1040 WIEN
AUSTRIA

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 345—348

COMPATIBLE FUZZY RELATIONS AND GROUPS

L. FILEP and I. GY. MAURER

Abstract

Let $(A, \ .)$ be a group, f a fuzzy set of A and R a fuzzy relation on f. In this paper the effect of the compatibility of R with $(A, \ .)$ to f and to the properties of R are examined.

1. Preliminaries

A mapping f from a non-empty set A to the real closed interval $[0, 1]$ is called a fuzzy subset, or a *fuzzy set* of A [5]. Denote $F(A)$ the set of all fuzzy subsets of A.

Let $f, g \in F(A)$, then the following concepts are well-known (see [2] and [3]): $f \subseteq g \Leftrightarrow f(x) \leqq g(x)$ (order), $(f \cap g)(x) = \min(f(x), g(x))$ (meet), $(f \cup g)(x) = \max(f(x), g(x))$ (union), $(f \times g)(x, y) = \min(f(x), g(y))$ (Cartesian product) for all $x, y \in A$;

An $R \subseteq f \times f \in F(A \times A)$ is called a fuzzy relation on f. It is said to be reflexive, if $R(x, x) = f(x) (\forall x \in A)$, symmetric, if $R(x, y) = R(y, x) (\forall x, y \in A)$, transitive, if $R(x, y) \geqq \min(R(x, z), R(z, y))$ $(\forall x, y, z \in A)$.

If $(A, \cdot)$ is a groupoid, then f will be called a fuzzy subgroupoid, or fuzzy groupoid, if $f(x \cdot y) \geqq \min(f(x), f(y))$ $(\forall x, y \in A)$. If $(A, \cdot)$ is a group, then the fuzzy groupoid f is said to be a fuzzy group on $(A, \cdot)$, if $f(x^{-1}) = f(x) (\forall x \in A)$.

According to the concept of the fuzzy congruence on an algebra [1], suggested by [4], we define the following concept for a group A, regarded as a universal algebra $(A, \cdot, {}^{-1}, e)$:

DEFINITION 1.1. Let A be a group and let R be a fuzzy relation on $f \in F(A)$. We say that R is

(i) *compatible with the binary operation* if

$$R(x_1 y_1, x_2 y_2) \geqq \min(R(x_1, y_1), R(x_2, y_2) \ (\forall x_1, y_1, x_2, y_2 \in A));$$

(ii) *compatible with the unary operation* $^{-1}$ if

$$R(x^{-1}, y^{-1}) \geqq R(x, y) \ (\forall x, y \in A);$$

(iii) *compatible with the nullary operation* e if

$$R(e, e) \geqq R(x, y) \ (\forall x, y \in A);$$

Research (partially) supported by Hungarian National Foundation for Scientific Research Grant nr. 1813.

1980 *Mathematics Subject Classification* (1985 Revision). Primary 03E72; Secondary 20N99.
Key words and phrases. Fuzzy group, compatible fuzzy relation.

(iv) a *compatible fuzzy relation on the group* A if it is compatible with every group operation.

We remark that if A is only a groupoid, then the compatibility of R means simply (i).

PROPOSITION 1.1. (ii)$\Leftrightarrow R(x^{-1}, y^{-1}) = R(x, y)(\forall x, y \in A)$.

PROOF. It follows from (ii) that $R(x, y) = R((x^{-1})^{-1}, (y^{-1})^{-1}) \geqq R(x^{-1}, y^{-1})$. Comparing this inequality with (ii) we get $R(x^{-1}, y^{-1}) = R(x, y)$. The converse statement is obvious. □

PROPOSITION 1.2. (i) *and* (ii)$\Rightarrow$(iii).

PROOF. Using (i) and Proposition 1.1 we get $R(e, e) = R(xx^{-1}, yy^{-1}) \geqq \geqq \min(R(x, y), R(x^{-1}, y^{-1})) = R(x, y)$ for all $x, y \in A$. □

On the basis of these propositions the notion of the compatibility can be characterized as follows:

THEOREM 1.1. *A fuzzy relation R on $f \in F(A)$ is a compatible fuzzy relation on the group A, iff*

1° $R(x_1 y_1, x_2 y_2) \geqq \min(R(x_1, y_1), R(x_2, y_2))$ $(\forall x_1, y_1, x_2, y_2 \in A)$;

2° $R(x^{-1}, y^{-1}) = R(x, y)$ $(\forall x, y \in A)$.

Further on we use frequently the following simple statement:

LEMMA 1.1. *If the fuzzy relation R on the fuzzy set $f \in F(A)$ is reflexive, then $R(x, x) \geqq R(x, y)$ and $R(y, y) \geqq R(x, y)$ for all $x, y \in A$.*

PROOF. $R(x, y) \leqq \min(f(x), f(y)) = \min(R(x, x), f(y)) \leqq R(x, x)$, and analogously $R(x, y) \leqq R(y, y)$. □

2. Results

Here we shall examine the effect of the compatibility defined above to the fuzzy set $f \in F(A)$ and to the properties of the fuzzy relation R on f.

THEOREM 2.1. *Let $(A, \cdot, ^{-1}, e)$ be a group and let R be a reflexive relation on the fuzzy set $f \in F(A)$. If R is compatible*

(a) *with the binary operation on A, then f is a fuzzy groupoid;*

(b) *with the unary operation on A, then $f(x^{-1}) = f(x)(\forall x \in A)$, and conversely;*

(c) *with both the binary and unary operations on A, then f is a fuzzy group.*

PROOF. (a) $R(x, x) = f(x)$, $R(y, y) = f(y) \Rightarrow R(xy, xy) \geqq \min(f(x), f(y))$. But $R(xy, xy) = f(xy)$. Therefore $f(xy) \geqq \min(f(x), f(y))$. This means that f is a fuzzy groupoid.

(b) $f(x^{-1}) = R(x^{-1}, x^{-1}) = R(x, x) = f(x)$

and

$$R(x^{-1}, x^{-1}) = f(x^{-1}) = f(x) = R(x, x).$$

(c) This statement is a consequence of (a) and (b). □

THEOREM 2.2. *Let A be a group and let R be a fuzzy relation on the fuzzy group $f\in F(A)$. If R is reflexive and compatible with the binary operation on A, then*
(a) *R is transitive;*
(b) *R is symmetric iff R is compatible with the unary operation on A.*

PROOF. (a) Let x, y, z be arbitrary elements of A. Using reflexivity of R and the fact that f is a fuzzy group, we can write

$$(*)\qquad R(x^{-1}, x^{-1}) = f(x^{-1}) = f(x) = R(x, x) \;(\forall x\in A).$$

Using also Lemma 1.1, it follows that

$$R(e, zx^{-1}) = R(xx^{-1}, zx^{-1}) \geqq \min\big(R(x, z), R(x^{-1}, x^{-1})\big) = \\ = \min\big(R(x, z), R(x, x)\big) = R(x, z),$$

and analogously

$$R(e, yz^{-1}) \geqq R(z, y).$$

Moreover, from these inequalities we get

$$R(e, yx^{-1}) = R\big(ee, yz^{-1}(zx^{-1})\big) \geqq \min\big(R(e, yz^{-1}), R(e, zx^{-1})\big) \geqq \\ \geqq \min\big(R(z, y), R(x, z)\big).$$

Likewise applying Lemma 1.1, it follows that

$$R(x, y) = R\big(ex, (yx^{-1})x\big) \geqq \min\big(R(e, yx^{-1}), R(x, x)\big) \geqq \\ \geqq \min\big(R(x, z), R(z, y)\big).$$

This means, that R is transitive.

(b) First let R be symmetric. Using the compatibility of R with the binary operation on A, as well as $(*)$ and Lemma 1.1, we get

$$R(e, x^{-1}y) = R(x^{-1}x, x^{-1}y) \geqq \min\big(R(x^{-1}, x^{-1}), R(x, y)\big) = \\ = \min\big(R(x, x), R(x, y)\big) = R(x, y).$$

It follows, using also the symmetry of R, that

$$R(x^{-1}, y^{-1}) = R(y^{-1}, x^{-1}) = R\big(ey^{-1}, (x^{-1}y)y^{-1}\big) \geqq \\ \geqq \min\big(R(e, x^{-1}y), R(y^{-1}, y^{-1})\big) = \\ = \min\big(R(e, x^{-1}y), R(y, y)\big) \geqq \min\big(R(x, y), R(y, y)\big) = R(x, y).$$

Thus $R(x^{-1}, y^{-1})\geqq R(x, y)$, which means that R is compatible with the unary operation on A.

To prove the converse statement, we use the inequality $R(e, y^{-1}x)\geqq R(y, x)$, which can be proved as the similar inequalities in part (a) of the proof. It follows, using also the compatibility of R with the unary operation on A, $(*)$ and Lemma 1.1, that

$$R(x, y) = R(x^{-1}, y^{-1}) = R\big(ex^{-1}, (y^{-1}x)x^{-1}\big) \geqq \\ \geqq \min\big(R(e, y^{-1}x), R(x^{-1}, x^{-1})\big) = \\ = \min\big(R(e, y^{-1}x), R(x, x)\big) \geqq \min\big(R(y, x), R(x, x)\big) = R(y, x).$$

Thus $R(x, y) \geqq R(y, x)$. Interchanging x and y we get similarly that $R(y, x) \geqq R(x, y)$. It follows that $R(x, y)=R(y, x)$, i.e. R is symmetric. □

COROLLARY. *For a torsion group A, from the conditions of Theorem 2.2 it follows that f is a fuzzy congruence on f.*

PROOF. Let x and y be two arbitrary elements of A, and denote m the least common multiple of the orders of x and y. Then $x^m=e$, $y^m=e$, and consequently $x^{-1}=x^{m-1}$, $y^{-1}=y^{m-1}$. Using the compatibility of R with the binary operation on A, we conclude by recursion that

$$R(x^{-1}, y^{-1}) = R(x^{m-1}, y^{m-1}) = R(x^{m-2}x, y^{m-2}y) \geqq$$
$$\geqq \min\left(R(x^{m-2}, y^{m-2}), R(x, y)\right) \geqq \ldots \geqq$$
$$\geqq \min\left(R(x, y), R(x, y), \ldots, R(x, y)\right) = R(x, y),$$

i.e. R is compatible with the unary operation on A. Then our statement follows from Theorem 2.2 (b). □

REFERENCES

[1] FILEP, L. and MAURER, I. GY., Fuzzy congruences and compatible fuzzy partitions, *Fuzzy Sets and Systems* **29** (1989), 357—361.
[2] NEGOITA, C. V. and RALESCU, D. A., *Applications of fuzzy sets to systems analysis*, ISR-Interdisciplinary Systems Research, [3] Vol. 11, Birkhäuser-Verlag, Basel—Stuttgart, 1975. *MR* **58** #9442b.
[3] ROSENFELD, A., Fuzzy groups, *J. Math. Anal. Appl.* **35** (1971), 512—517.
[4] VOJVODIC, G. and ŠEŠELJA, B., The structure of fuzzy equivalence relations and fuzzy congruence relations, *Mat. Vesnik* **1 (14) (29)** (1977), 147—152 (in Serbian). *MR* **57** #5762.
[5] ZADEH, L. A., Fuzzy sets, *Information and Control* **8** (1965), 338—353.

(Received May 1, 1987)

BESSENYEI GYÖRGY TANÁRKÉPZŐ FŐISKOLA
MATEMATIKA TANSZÉK
SÓSTÓI ÚT 31/B
H—4400 NYÍREGYHÁZA

NEHÉZIPARI MŰSZAKI EGYETEM
MATEMATIKAI INTÉZET
H—3515 MISKOLC—EGYETEMVÁROS
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 349—354

SOME CONDITIONS FOR A SELF-INJECTIVE RING TO BE QUASI-FROBENIUS

NGUYEN V. DUNG

1. Introduction

Since several years, various conditions for a self-injective ring to be quasi-Frobenius have been given by many authors. A well-known result due to C. Faith [3] asserts that a right (or left) self-injective ring with the ACC on annihilator right ideals is a quasi-Frobenius ring. J. Lawrence [7] has shown that any countable right self-injective ring is quasi-Frobenius. As the main result of [1], E. P. Armendariz has proved that right or left self-injective rings satisfying the minimum condition on essential right ideals are quasi-Frobenius.

This paper is essentially motivated by [1]. We prove here that a right self-injective ring R is quasi-Frobenius if R satisfies the right restricted minimum condition (Theorem 3.2), or if R satisfies the maximum condition on essential right ideals (Theorem 3.6). Self-injective rings with some other restricted forms of the maximum condition are also investigated. The method of our proofs differs substantially from that of Armendariz in [1], and can be used to simplify a part of the proof of [1, Theorem]. Some results of Yue Chi Minh [10] are also considerably improved.

2. Definitions and notation

Throughout this paper we consider associative rings with identity element and unitary modules. We will write M_R (resp. ${}_RM$) to indicate that M is being considered as a right (resp. left) R-module.

Let R be a ring and let M be a right R-module. The socle of M will be denoted by soc (M). Saying that M has finite Goldie dimension means that M has no infinite direct sums of non-zero submodules. A submodule K of M is essential if $K \cap L \neq (0)$ for every non-zero submodule L of M. A right ideal of R is essential if it is essential as a submodule of R_R. The ring R is said to satisfy the right restricted minimum condition (right RMC) if the right R-module R/E is Artinian for every essential right ideal E of R.

The Jacobson radical of a ring R will be denoted by $J(R)$, the prime radical by $N(R)$ and the right singular ideal by $Z(R)$. For any subset X of R, $r(X)$ (resp. $l(X)$) represents the right (resp. left) annihilator of X in R. The abbreviations ACC and DCC stand for the ascending chain condition and the descending chain condition, respectively.

1980 *Mathematics Subject Classification* (1985 Revision). Primary 16A36, 16A52; Secondary 16A34.

Key words and phrases. Self-injective ring, quasi-Frobenius ring, injective module, essential submodule.

3. Some conditions for a self-injective ring to be quasi-Frobenius

We start with a lemma which will be used repeatedly in this paper.

LEMMA 3.1. *Let M be a finitely generated injective right R-module. Suppose that for every essential submodule K of M, M/K has a finitely generated socle (possible zero). Then M and all factor modules of M have finite Goldie dimension.*

PROOF. Let $C\subseteq D$ be submodules of M such that C is an essential submodule of D. By Zorn's Lemma there exists a submodule L of M which is maximal with respect to the property that $C\cap L=(0)$. Then $D\cap L=(0)$ and $C\oplus L$ is an essential submodule of M, hence $M/(C\oplus L)$ has a finitely generated socle. Since D/C is R-isomorphic to $(D\oplus L)/(C\oplus L)$, it follows that D/C has a finitely generated socle.

Suppose that there exists an infinite direct sum $A=A_1\oplus A_2\oplus A_3\ldots$ of non-zero submodules A_i in M. Then we have an infinite direct sum $A=B_1\oplus B_2\oplus B_3, \ldots,$ where each B_i is an infinite direct sum $B_i=A_{i_1}\oplus A_{i_2}\oplus A_{i_3}\ldots$. We denote by $E(A)$ the injective hull of A in M and by $E(B_i)$ the injective hull of B_i in $E(A)$. It is clear that $E(B_1)\oplus E(B_2)\oplus E(B_3)\ldots$ is an infinite direct sum. Because M is finitely generated, $E(B_i)$ is also finitely generated, hence $E(B_i)\neq B_i$ for each i. So for each i we can choose an element $x_i\in E(B_i)$ such that $x_i\notin B_i$. There exists a submodule $C_i\subseteq E(B_i)$ which is maximal with respect to the property that $B_i\subseteq C_i$ and $x_i\notin C_i$. Then $E(B_i)/C_i$ contains a simple submodule, therefore $(E(B_1)\oplus E(B_2)\oplus\ldots)/(C_1\oplus C_2\oplus\ldots)$ has an infinitely generated socle. But clearly $C_1\oplus C_2\oplus\ldots$ is essential in $E(B_1)\oplus \oplus E(B_2)\oplus\ldots$, hence $(E(B_1)\oplus E(B_2)\oplus\ldots)/(C_1\oplus C_2\oplus\ldots)$ must have a finitely generated socle, a contradiction. Thus we have shown that M has finite Goldie dimension.

Let K be any submodule of M and let $E(K)$ be the injective hull of K in M. Then $M=E(K)\oplus L$ for a submodule L of M. We have $M/K\cong(E(K)/K)\oplus L$, so M/K has a finitely generated socle. Therefore each factor module of M has a finitely generated socle. Repeating the above argument, we can show that M/K has finite Goldie dimension, because otherwise a factor module of M/K would have an infinitely generated socle.

The proof of Lemma 3.1 is complete.

THEOREM 3.2. *Let R be a right self-injective ring satisfying the right restricted minimum condition. Then R is a quasi-Frobenius ring.*

PROOF. Let J be the Jacobson radical of R. Then $J=Z(R)$ and R/J is a von Neumann regular right self-injective ring (see, e.g. [4, Corollary 19.28]). By Lemma 3.1, R/J has finite right Goldie dimension. Thus R/J is semisimple Artinian. Let A be a non-zero right ideal of R. If $A\cap J=(0)$, then A is R-isomorphic to a right ideal of R/J, hence A is Artinian as a right R-module. Now suppose that $A\cap J\neq(0)$, then there exists a non-zero element $x\in A\cap J$. We have $xR\cong R/r(x)$. Because $J=Z(R)$, $r(x)$ is an essential right ideal of R. Since R satisfies the right restricted minimum condition, $R/r(x)$ is Artinian, hence xR is Artinian. Therefore every non-zero right ideal A of R contains a non-zero minimal right ideal. It follows that the right socle S of R is an essential right ideal, thus R/S is Artinian. By Lemma 3.1, R has finite right Goldie dimension, hence S is finitely generated as a right R-module. Thus R is right Artinian, hence a quasi-Frobenius ring.

COROLLARY 3.3 (Armendariz [1]). *Let R be a right self-injective ring satisfying the minimum condition on essential right ideals. Then R is a quasi-Frobenius ring.*

REMARK. A ring R satisfies the minimum condition on essential right ideals if and only if $R/S(R)$ is Artinian as a right R-module, where $S(R)$ is the right socle of R (see [1, Proposition 1.1]). If R satisfies the minimum condition on essential right ideals, then R satisfies the right RMC, obviously. However, the converse is not true, in general. J. Cozzens [2] has constructed an example of a simple right Ore domain R such that R/K is semisimple for every non-zero right ideal K of R, but R is not Artinian.

Theorem 3.2 naturally raises the following question.

QUESTION 1. Let R be a left self-injective ring satisfying the right restricted minimum condition. Is R a quasi-Frobenius ring?

Next we shall examine rings with the maximum condition on essential right ideals. The proof of the following result is a modification of the proof of [5, Proposition 3.6].

LEMMA 3.4. *An R-module M satisfies the maximum condition on essential submodules if and only if $M/\mathrm{soc}(M)$ is Noetherian.*

PROOF. It is well-known that $S=\mathrm{soc}(M)$ is the intersection of all essential submodules of M. Hence if M/S is Noetherian, then M satisfies the maximum condition on essential submodules. Conversely, assume that M satisfies the maximum condition on essential submodules. Then M/E is Noetherian for every essential submodule E of M. Let $A\subseteq B$ be submodules of M such that A is an essential submodule of B. There exists a submodule C of M such that $A\cap C=(0)$ and $A\oplus C$ is an essential submodule of M. Then $B\cap C=(0)$ and $B/A\cong(B\oplus C)/(A\oplus C)$, hence B/A is Noetherian since $M/(A\oplus C)$ is Noetherian.

Let L be any submodule of M with $L\cap S=(0)$. We will show that L is finitely generated. Suppose that there exists an infinite direct sum $X=X_1\oplus X_2\oplus X_3\dots$ of non-zero submodules X_i contained in L. Then each X_i has a proper essential submodule Y_i. It follows that $Y=Y_1\oplus Y_2\oplus\dots$ is an essential submodule of X and hence X/Y is Noetherian. But $X/Y\cong(X_1/Y_1)\oplus(X_2/Y_2)\oplus\dots$, a contradiction. Thus L has finite Goldie dimension. We can choose a maximal independent family $\{E_j\}$ of non-zero cyclic submodules of L. Then $E=\sum_j^{\oplus} E_j$ is a finitely generated essential submodule of L. Since L/E is Noetherian it follows that L is finitely generated.

Let D be any submodule of M containing S. There exists a submodule L of D such that $L\cap S=(0)$ and $L-S$ is an essential submodule of D. Then $D/(L\oplus S)$ is Noetherian, and L is finitely generated by the above argument. It implies that D/S is finitely generated. Hence M/S is a Noetherian module.

THEOREM 3.5. *Let M be a finitely generated injective right R-module satisfying the maximum condition on essential submodules. Then M is Noetherian.*

PROOF. By Lemma 3.1, M has finite Goldie dimension, hence $\mathrm{soc}(M)$ is finitely generated. Since M satisfies the maximum condition on essential submodules, from Lemma 3.4 it follows that $M/\mathrm{soc}(M)$ is Noetherian. Thus M is Noetherian.

As a special case of Theorem 3.5 we have

THEOREM 3.6. *Let R be a right self-injective ring with the maximum condition on essential right ideals. Then R is a quasi-Frobenius ring.*

REMARK. A ring R is called an ETLA-ring if all essential left ideals and ideals of R are annihilator left ideals (see [10]). In [10, Theorem 2.3] Yue Chi Minh has proved the equivalence of the following conditions: (1) R is a quasi-Frobenius ring; (2) R is a left and right self-injective ETLA-ring with maximum condition on essential left ideals and ideals; (3) R is a left self-injective right cogenerating ETLA-ring with maximum condition on essential left ideals and ideals.

It is clear that this result is an immediate consequence of Theorem 3.6.

In connection with Theorem 3.6 we have

QUESTION 2. Let R be a left self-injective ring satisfying the maximum condition on essential right ideals. Is R a quasi-Frobenius ring?

We have not been able to settle this question. However, the result below presents some strong evidence for its truth.

PROPOSITION 3.7. *Let R be a left self-injective ring satisfying the maximum condition on essential right ideals. Then R is a semiperfect ring with nilpotent prime radical $N(R)$ such that $R/N(R)$ is right Noetherian. Furthermore, if $N(R)=J(R)$, then R is a quasi-Frobenius ring.*

PROOF. Briefly we set $J=J(R)$ and $N=N(R)$. It is easy to prove that every factor ring of R also satisfies the maximum condition on essential right ideals. By [4, Corollary 19.28], $\bar{R}=R/J$ is a von Neumann regular left self-injective ring. Let $\bar{S}$ be the right socle of $\bar{R}$, then $\bar{S}$ is also the left socle of $\bar{R}$. By Lemma 3.4, $\bar{R}/\bar{S}$ is right Noetherian, hence a semisimple Artinian ring. Thus $\bar{R}/\bar{S}$ is left Artinian. From Lemma 3.1 it follows that $\bar{R}$ has finite left Goldie dimension, hence $\bar{R}$ is semisimple Artinian. Since R is left self-injective, idempotents in R/J can be lifted to R (see [4, Theorem 19.27]), thus R is a semi-perfect ring.

Let S be the right socle of R. Then $(N+S)/S\subseteq N(R/S)$. Since R/S is right Noetherian, $N(R/S)$ is nilpotent. Hence $N^k\subseteq S$ for some integer $k\geqq 1$. But $SN=(0)$, it follows that $N^{k+1}=(0)$. Now we consider the ring $R^*=R/N$ with the right socle S^*. Since R is semi-perfect, it is clear that R^* is a semiprime ring which does not contain an infinite system of orthogonal idempotents. From this it follows that S^* is a finitely generated right ideal of R^* (see e.g. [6]). By Lemma 3.4, R^*/S^* is right Noetherian, hence R^* is right Noetherian.

Suppose that $N=J$. Then J is nilpotent and R/J is a semisimple Artinian ring. Now we can use the argument of Armendariz in [1, p. 307]. Since R is a semi-primary ring, R has an essential left socle, hence R is a left PF-ring (see e.g. [4, Proposition 24.32]). It follows that $J=lr(J)$ (see [4, 24.3(f)]). Since R/J is Artinian, there exists a finitely generated right ideal A contained in $r(J)$ such that $J=l(A)$. Because R is left self-injective, A is a right annihilator, so we have $A=rl(A)=r(J)$. Note that $r(J)$ coincides with the right socle S of R. Thus S is a finitely generated right ideal of R. Since R/S is right Noetherian, it implies that R is right Noetherian, hence a quasi-Frobenius ring.

[1, Theorem] and our previous results suggest the following

QUESTION 3. Let R be a right (or left) self-injective ring with right socle S such that the ring R/S satisfies the ACC on annihilator right ideals. Is R a quasi-Frobenius ring?

The result below shows that it is true for semiprime rings. On the other hand, it extends [1, Lemma 2.3].

THEOREM 3.8. *Let R be a semiprime ring with the right socle S such that the ring R/S satisfies the ACC on annihilator right ideals. If R is left or right self-injective, then R is a semisimple Artinian ring.*

PROOF. Suppose first that R is left self-injective. Since R is a semiprime ring, S coincides with the left socle of R, and every minimal left ideal of R is generated by an idempotent. Let J be the Jacobson radical of R. It is well-known that J does not contain non-zero idempotents, hence $S\cap J=(0)$. Let H be the injective hull of S in ${}_RR$. Then $H\cap J=(0)$ and $H=Re$ for some idempotent $e\in R$. Note that $(1-e)S\cdot\cdot(1-e)S=(0)$, hence $(1-e)S=(0)$ because R is semiprime, and so $R(1-e)\subseteq l(S)$. On the other hand, we have $l(S)\cap S=(0)$, hence $l(S)\cap Re=(0)$. Thus $l(S)==R(1-e)$. From this it follows that $H=Re$ is a two-sided ideal and e is a central idempotent of R. Since R is left self-injective, R/J is a von Neumann regular ring, thus H is also a regular ring because $H\cap J=(0)$. Therefore R has a ring decomposition $R=H\oplus F$, where H is a von Neumann regular left self-injective ring, and F is a semiprime left self-injective ring. We note that S is the right (and left) socle of H and $R/S\cong(H/S)\oplus F$. Since R/S satisfies the ACC for annihilator right ideals, the rings F and H/S also satisfy the ACC for annihilator right ideals. Then F is a semisimple Artinian ring by Faith [3]. In a regular ring every finitely generated right ideal is an annihilator right ideal, thus H/S satisfies the ACC on finitely generated right ideals. This implies that H/S is right Noetherian, hence semisimple Artinian. Thus H/S is left Artinian. By Lemma 3.1 it implies that H has finite left Goldie dimension, therefore H is semisimple Artinian. We conclude that R is a semisimple Artinian ring.

The case if R is right self-injective can be proved similarly.

A well-known result of B. Osofsky [8] asserts that a ring R is semisimple Artinian if and only if every cyclic right R-module is injective. Following P. F. Smith [9], a ring R is said to satisfy the right restricted injective condition (shortly right RIC-ring) if R/K is an injective right R-module for every essential right ideal K of R. The next proposition is a slight generalization of Osofsky's result.

PROPOSITION 3.9. *Let R be a right self-injective right RIC-ring. Then R is a semisimple Artinian ring.*

PROOF. Let A be any right ideal of R and $E(A)$ be the injective hull of A in R_R. Then $E(A)=eR$ for some idempotent e in R. It is clear that $A\oplus(1-e)R$ is an essential right ideal of R, hence $R/(A\oplus(1-e)R)\cong eR/A$ is an injective right R-module. Then we have $R/A=(eR\oplus(1-e)R)/A\cong(eR/A)\oplus(1-e)R$, and so R/A is an injective right R-module. Therefore every cyclic right R-module is injective and R is a semisimple Artinian ring by Osofsky [8].

Proposition 3.9 raises the following question:

QUESTION 4. Let R be a left self-injective right RIC-ring. Is R a semisimple Artinian ring?

REMARK. It is clear that a ring R is right RIC if and only if every finitely generated singular right R-module is injective. Following K. R. Goodearl [5], a ring R is called a right SI-ring if every singular right R-module is injective. We do not know any example of a right RIC-ring which is not a right SI-ring. Now let R be a left self-injective right SI-ring. By [5, Theorem 3.11] there is a ring decomposition $R = K \oplus R_1 \oplus \ldots \oplus R_n$ such that $K/\mathrm{soc}\,(K_K)$ is semisimple Artinian and R_i is right Noetherian for each i. Then by [1, Theorem] R is quasi-Frobenius, hence semisimple Artinian. This suggests an affirmative answer to Question 4.

ADDED IN PROOF. Dinh Van Huynh, R. Wisbauer and the present author have recently obtained a positive answer to Question 2. This result is contained in the forthcoming paper „Quasi-injective modules with ACC or DCC on essential submodules", *Arch. Math. (Basel)*.

REFERENCES

[1] ARMENDARIZ, E. P., Rings with DCC on essential left ideals, *Comm. Algebra* **8** (1980), 299—308. *MR* **81c**: 16021.

[2] COZZENS, J. H., Homological properties of the ring of differential polynomials, *Bull. Amer. Math. Soc.* **76** (1970), 75—79. *MR* **41** # 3531.

[3] FAITH, C., Rings with ascending condition on annihilators, *Nagoya Math. J.* **27** (1966), 179—191. *MR* **33** # 1328.

[4] FAITH, C., *Algebra II: Ring Theory*, Grundlehren der mathematischen Wissenschaften, Vol. 191, Springer-Verlag, Berlin—Heidelberg—New York, 1976. *MR* **55** # 383.

[5] GOODEARL, K. R., *Singular torsion and the splitting properties*, Memoirs of the Amer. Math. Soc. No. 124, American Mathematical Society, Providence, R. I., 1972. *MR* **49** # 5090.

[6] GORDON, R., Rings in which minimal left ideals are projective, *Pacific J. Math.* **31** (1969), 679—692. *MR* **42** # 314.

[7] LAWRENCE, J., A countable self-injective ring is quasi-Frobenius, *Proc. Amer. Math. Soc.* **65** (1977), 217—220. *MR* **56** # 414.

[8] OSOFSKY, B. L., Non-injective cyclic modules, *Proc. Amer. Math. Soc.* **19** (1968), 1383—1384.

[9] SMITH, P. F., Rings characterized by their cyclic modules, *Canad. J. Math.* **31** (1979), 93—111. *MR* **80d**: 16018.

[10] YUE CHI MING, R., On regular ideals and P-injectivity, *Rendiconti Seminario Fac. Scienze Univ. Cagliari* **54** (1984), 25—34.

(Received May 2, 1987)

INSTITUTE OF MATHEMATICS
P.O. BOX 631 BỜ HỒ
HANOI
VIETNAM

CONTENTS

HU ISSN 0081-6906

Index: 26.748

Studia Scientiarum Mathematicarum Hungarica

VOLUME 24
NUMBER 4
1989

AKADÉMIAI KIADÓ, BUDAPEST

STUDIA SCIENTIARUM MATHEMATICARUM HUNGARICA

A QUARTERLY OF THE HUNGARIAN ACADEMY OF SCIENCES

Studia Scientiarum Mathematicarum Hungarica publishes original papers on mathematics mainly in English, but also in German, French and Russian.

Studia Scientiarum Mathematicarum Hungarica is published in yearly volumes of four issues (mostly double numbers published semiannually) by

AKADÉMIAI KIADÓ
Publishing House of the Hungarian Academy of Sciences
H—1117 Budapest, Frielle Kornélia u. 19—35.

Manuscripts and editorial correspondence should be addressed to

J. Merza
Managing Editor
P.O. Box 127
H—1364 Budapest

Subscription information

Orders should be addressed to

KULTURA Foreign Trading Company
P.O. Box 149
H—1389 Budapest

or to its representatives abroad.

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 355—364

ON THE NUMBER OF DIFFERENT PATTERNS PRECEDING A GIVEN ONE

TAMÁS F. MÓRI

1. Introduction

Let H be a finite alphabet of size d and let H^n denote the set of length n words over H. Consider an infinite sequence of random letters, i.e. let $X_1, X_2, \dots$ be i.i.d. random variables taking on the elements of H each with probability $1/d$. For any word $A \in H^n$ let

$$T(A) = \inf\{m\colon (X_{m-n+1}X_{m-n+2}\dots X_m) \equiv A\},$$

this is the waiting time till A is observed as a run in the process X_i.

Paul Erdős posed the following problem. Find the asymptotics as $k \to \infty$ of the maximum number M_k for which every word of that length has been observed as a run during the first k experiments. That is,

$$M_k = \max\{n\colon T(A) \leqq k \text{ for every } A \in H^n\}.$$

The answer is given in [7], where the sequence M_k is shown to be asymptotically quasideterministic; with probability 1

$$\left[\frac{1}{\log d}(\log k - \log\log k - \varepsilon)\right] \leqq M_k \leqq \left[\frac{1}{\log d}(\log k - \log\log k + \varepsilon)\right]$$

holds for any $\varepsilon > 0$ if k is large enough.

Using the Conway algorithm (see [4] for an elegant proof) one can easily show that pure runs (=homogeneous patterns) need the longest time to occur. Let A be a pure run and $B \neq A$ an arbitrary word of the same length, then $\mathsf{P}(T(B) \leqq T(A)) \geqq 1/2$ and also $\mathsf{E}(T(B)) \leqq \mathsf{E}(T(A))$. This makes it likely that pure runs are more frequently the last of H^n to appear than any other word. Fixing a letter α let us denote by N_k the length of the longest pure α-run observed in the first k experiments. It is well-known [2] that with probability 1

$$\left[\frac{1}{\log d}(\log k - \log\log\log k) - 1 - \varepsilon\right] \leqq N_k$$

Research supported by the Hungarian National Foundation for Scientific Research Grant No. 1808.

1980 *Mathematics Subject Classification* (1985 Revision). Primary 60F15; Secondary 60C05.
Key words and phrases. Run, sequence pattern, waiting time, independence principle, Borel—Cantelli lemma, Conway algorithm.

holds for large values of k. Comparing this with (1), we immediately see that for large n a pure run can never be the last of H^n to appear. Erdős asked [1] if it can happen infinitely often that when a pure run is first observed, almost every word of the same length has already appeared. Here we answer this question in the affirmative, describing the limit behaviour of the number of different words observed before a given word appears.

2. Applying the independence principle

Let us give the problem a generalized formulation. Suppose we are given a subset of words $H_n \subset H^n$ of size $|H_n| = N$. Further, fix $B \in H^n$ and denote by R_B the proportion of words in H_n which had not been seen before B was first observed. Finally, let $\beta = d^n/\mathsf{E}(T(B))$, this quantity can easily be calculated by the Conway algorithm. For any two words $A, B \in H^n$ define the *leading number* of A over B as $A*B = \sum_{i=1}^{n} \varepsilon_i d^i$, where $\varepsilon_i = 1$ if the last i letters of A coincide with the first i letters of B, otherwise $\varepsilon_i = 0$. Then $\mathsf{E}(T(B)) = B*B$. Obviously, $1 - \frac{1}{d} < \beta \leq 1$.

In [6] the following approximate description of the joint asymptotic behaviour of the waiting times $\{T(A)\colon A \in H^n\}$ is given. In some cases they can be regarded as independent, exponentially distributed random variables with common expectation d^n. By using this approximation one can often guess the limit distribution of certain functionals of the waiting times. Then an exact proof can be given upon the basis of [5], which paper thus gives a precise mathematical justification to the above heuristic method called *independence principle*.

Let us see how the independence principle applies to our present problem. Clearly,

$$R_B = \frac{1}{N} \sum_{A \in H_n} I(T(A) > T(B)),$$

where I stands for the indicator function. Even given $T(B) = T$, the events $\{T(A) > T(B)\}$ are "conditionally independent" with the same conditional probability $p = \exp(-d^{-n}T)$ assigned to each. If N is large enough, the law of large numbers implies that the conditional distribution of R_B is concentrated onto a small neighbourhood of p. Hence

$$R_B^{\beta} \approx \exp\left(-\frac{T(B)}{\mathsf{E}(T(B))}\right).$$

The distribution of the right-hand side is approximately uniform $U(0, 1)$, that is, R_B has a Beta $(\beta, 1)$ distribution in the limit. A detailed, precise reasoning is given in Section 4.

3. Results

Our first assertion deals with the accuracy achieved when R_B is approximated by $S_B = \exp(-d^{-n}T(B))$ and by a Beta $(\beta, 1)$ variable, resp.

THEOREM 1.

(a)
$$\mathsf{E}\,|R_B - S_B| \leqq 14N^{-1/6}.$$

(b)
$$\sup_{0\leqq t\leqq 1} |\mathsf{P}(R_B \leqq t) - t^\beta| \leqq 8N^{-(\beta/2(2+\beta))}.$$

Note that the exponent of N in part (b) lies between $-\frac{1}{6}$ and $-\frac{d-1}{6d-2}$.

Now consider a sequence of words $B_n \in H^n$ and sets $H_n \subset H^n$, $n=1, 2, \ldots$. Our Theorem 1(b) then provides a weak limit theorem for $R_n = R_{B_n}$, together with an estimation for the rate of convergence. Next we are interested in the a.s. asymptotic behaviour of the sequences $S_n = S_{B_n}$ and $R_n = R_{B_n}$.

THEOREM 2. (a) *Let (t_n), (u_n) be sequences of real numbers, $0 < t_n < 1$, $0 < u_n < 1$.*

If $\sum_{n=1}^{\infty} t_n^{\beta_n} < \infty$, then with probability 1 $S_n > t_n$ for every n large enough.

If $\sum_{n=1}^{\infty} t_n^{\beta_n} = \infty$, then with probability 1 $S_n \leqq t_n$ infinitely often.

If $\sum_{n=1}^{\infty} u_n < \infty$, then with probability 1 $S_n < 1 - u_n$ for every n large enough.

If $\sum_{n=1}^{\infty} u_n = \infty$, then with probability 1 $S_n \geqq 1 - u_n$ infinitely often.

(b) *Suppose*

$$\sum_{n=1}^{\infty} N_n^{-(\beta_n/6(1+\beta_n))} < \infty.$$

Then the characterization given in (a) *is also valid for R_n in place of S_n.*

Remembering the bounds of β_n we find that the condition in part (b) is met if

$$\sum_{n=1}^{\infty} N_n^{-1/18} < \infty.$$

As a consequence of Theorem 2 we can give a complete answer to the question of Erdős, by substituting $N_n = d^n$ and $\beta_n = \frac{1-d^{-1}}{1-d^{-n}}$ $\Big($the latter can be replaced equivalently by $1 - \frac{1}{d}\Big)$.

Especially, we obtain the following assertion. Let $\alpha \in H$ be fixed. Then it can happen with probability 1 for infinitely many n that when an α-run of length n is first observed, there are less than $n^{-d/(d-1)}d^n$ words of length n that have not been seen so far. On the other hand, with finitely many exceptions for n, one can always find at least $n^{-d/(d-1)-\varepsilon}d^n$ words that occur later than a pure α-run of length n.

Part (a) of Theorem 2 describes the a.s. variation of the sequence $T(B_n)$ for an arbitrary sequence of words $B_n \in H^n$. Such assertions were announced in [3], though with a small error, in the form of unnumbered text remarks without proofs. For the sake of completeness we do give a proof here in Section 5.

4. Proof of Theorem 1

We can suppose $N^{1/6} \geqq 14$, otherwise the assertion is trivial. Let $m = \left[\frac{1}{4} N^{1/6}\right]$, then $m \geqq 3$ and hence $m \geqq \frac{3}{16} N^{1/6}$. It is not so difficult to see that

$$(2) \qquad \begin{aligned} \mathsf{E}\,|R_B - S_B| \leqq \frac{2}{m} + \frac{1}{m} \sum_{i=1}^{m-2} \mathsf{P}\left(R_B < \frac{i}{m}, \frac{i+1}{m} \leqq S_B\right) + \\ + \frac{1}{m} \sum_{i=1}^{m-2} \mathsf{P}\left(S_B < \frac{i}{m}, \frac{i+1}{m} \leqq R_B\right). \end{aligned}$$

Here $S_B < \frac{i}{m}$ if and only if $T(B) > d^n \log \frac{m}{i}$. Introduce the random variables

$$R(t) = \frac{1}{N} \sum_{A \in H_n} I\left(T(A) > d^n \log \frac{1}{t}\right), \qquad 0 < t < 1.$$

Then

$$\mathsf{P}\left(R_B < \frac{i}{m}, \frac{i+1}{m} \leqq S_B\right) \leqq \mathsf{P}\left(R\left(\frac{i+1}{m}\right) < \frac{i}{m}\right)$$

and

$$\mathsf{P}\left(S_B < \frac{i}{m}, \frac{i+1}{m} \leqq R_B\right) \leqq \mathsf{P}\left(\frac{i+1}{m} \leqq R\left(\frac{i}{m}\right)\right).$$

These probabilities will be estimated by the help of the Chebyshev inequality. Let us first calculate the variance of $R(t)$. Clearly,

$$\begin{aligned} \operatorname{Var}(R(t)) = \frac{1}{N^2} \sum_{A_1, A_2 \in H_n} \Bigg(\mathsf{P}\left(T(A_1) > d^n \log \frac{1}{t}, T(A_2) > d^n \log \frac{1}{t}\right) - \\ - \mathsf{P}\left(T(A_1) > d^n \log \frac{1}{t}\right) \mathsf{P}\left(T(A_2) > d^n \log \frac{1}{t}\right)\Bigg) = \\ = \frac{1}{N^2} \sum_{A_1, A_2 \in H_n} \delta(A_1, A_2). \end{aligned}$$

Tail probabilities for distributions of waiting times can be estimated by using the results of [5]. Let $A_1, A_2, \ldots, A_r$ be different words of length n, and denote by Z

the waiting time till any of them appears: $Z=\min T(A_i)$. Introduce the notation $b=2rnd^{-n}$, $c=r(r-1)d^{-n}\max\limits_{i\neq j} A_i*A_j$. Then for every positive x we have

$$(3)\ \exp\left(-(1+b)\,x\sum_{i=1}^{r}\frac{1}{A_i*A_i}\right)\leqq \mathsf{P}(Z>x)\leqq(1+b)\exp\left(-(1-c)\,x\sum_{i=1}^{r}\frac{1}{A_i*A_i}\right),$$

provided $b<1/5$. This inequality immediately implies that

$$(1-b)\Big(\prod_{i=1}^{r}\mathsf{P}(T(A_i)>x)\Big)^{1+b}\leqq \mathsf{P}(Z>x)\leqq(1+b)\Big(\prod_{i=1}^{r}\mathsf{P}(T(A_i)>x)\Big)^{1-b-c},$$

from which one can easily derive the estimation

$$(4)\qquad -2b<\mathsf{P}(Z>x)-\prod_{i=1}^{r}\mathsf{P}(T(A_i)>x)\leqq 2b+c.$$

Let us consider the pairs $(A_1, A_2)\in H_n\times H_n$ for which $A_1*A_2\leqq N^{-1/2}d^n$, $A_2*A_1\leqq N^{-1/2}d^n$ (this also implies $A_1\neq A_2$). Applying (4) with $r=2$, $x=d^n\log\dfrac{1}{t}$, we obtain

$$|\delta(A_1, A_2)|=8\,nd^{-n}+2N^{-1/2}\leqq(8\,nd^{-n/2}+2)N^{-1/2}\leqq 10N^{-1/2}.$$

For all the other pairs

$$|\delta(A_1, A_2)|\leqq 1.$$

It is simple to prove that for any fixed $A_1\in H^n$ the number of words $A_2\in H^n$ for which either A_1*A_2 or A_2*A_1 exceeds $N^{-1/2}d^n$ is less than $8\sqrt{N}$ (see [8] for a short proof). Hence

$$(5)\qquad \operatorname{Var}(R(t))\leqq 18N^{-1/2}.$$

We will now turn our attention to the expectation of $R(t)$. Let us start from the formula

$$\mathsf{E}(R(t))=\frac{1}{N}\sum_{A\in H_n}\mathsf{P}\left(T(A)>d^n\log\frac{1}{t}\right),$$

then we can use (3) with $r=1$ to show that

$$(6)\qquad \left|\mathsf{P}(T(A)>x)-\exp\left(-\frac{x}{A*A}\right)\right|\leqq 2nd^{-n}\leqq 2N^{-1/2}.$$

Since

$$0\leqq\exp\left(-\frac{x}{A*A}\right)-\exp(-d^{-n}n)\leqq d^{-n}\,A*A-1,$$

it seems reasonable to classify the words $A\in H_n$ according that $A*A$ is less or greater than $d^n(1+N^{-1/2})$. For words with $A*A\leqq d^n(1+N^{-1/2})$ we have

$$\left|\mathsf{P}\left(T(A)>d^n\log\frac{1}{t}\right)-t\right|\leqq 3N^{-1/2},$$

while the number of other words can be majorized by $4\sqrt{N}$ (we refer to [8] again). Hence

$$\left|\mathsf{E}(R(t))-t\right| \leqq 7N^{-1/2}. \tag{7}$$

Here $7N^{-1/2}<\frac{1}{20}N^{-1/6}$. Now by the Chebyshev inequality

$$\mathsf{P}\left(R\left(\frac{i+1}{m}\right)<\frac{i}{m}\right) \leqq \mathsf{P}\left(\left|R\left(\frac{i+1}{m}\right)-\mathsf{E}\left(R\left(\frac{i+1}{m}\right)\right)\right|>\mathsf{E}\left(R\left(\frac{i+1}{m}\right)\right)-\frac{i}{m}\right) \leqq$$

$$\leqq \mathrm{Var}\left(R\left(\frac{i+1}{m}\right)\right)\Big/\left(\mathsf{E}\left(R\left(\frac{i+1}{m}\right)\right)-\frac{i}{m}\right)^2 \leqq$$

$$\leqq 18N^{-1/2}\left(\frac{79}{20}N^{-1/6}\right)^{-2} < \frac{3}{2}N^{-1/6},$$

and similarly,

$$\mathsf{P}\left(\frac{i+1}{m} \leqq R\left(\frac{i}{m}\right)\right) < \frac{3}{2}N^{-1/6}, \qquad i = 1, 2, \ldots, m-1.$$

Substituting these into (2) we obtain that

$$\mathsf{E}\,|R_B-S_B| \leqq \frac{2}{m}+3N^{-1/6} < 14N^{-1/6}.$$

In order to prove part (b) let $m=\left[\frac{1}{2}N^{1/2(2+\beta)}\right]$. We suppose $N^{\beta/2(2+\beta)}\geqq 8$, hence $m\geqq\frac{2}{5}N^{1/2(2+\beta)}$. If $0<t<1-\frac{1}{m}$,

$$\begin{aligned}\mathsf{P}(R_B \leqq t) &\leqq \mathsf{P}\left(S_B \leqq t+\frac{1}{m}\right)+\mathsf{P}\left(S_B > t+\frac{1}{m},\, R_B \leqq t\right) \leqq \\ &\leqq \mathsf{P}\left(S_B \leqq t+\frac{1}{m}\right)+\mathsf{P}\left(R\left(t+\frac{1}{m}\right) \leqq t\right).\end{aligned} \tag{8}$$

The first term on the right-hand side is estimated by (6):

$$\mathsf{P}\left(S_B \leqq t+\frac{1}{m}\right) = \mathsf{P}\left(T(B) > -d^n \log\left(t+\frac{1}{m}\right)\right) \leqq$$

$$\leqq \left(t+\frac{1}{m}\right)^{\beta}+2N^{-1/2} \leqq t^{\beta}+m^{-\beta}+2N^{-1/2}.$$

To the second term we apply the Chebyshev inequality. By (5) and (7)

$$\mathsf{P}\left(R\left(t+\frac{1}{m}\right)\leqq t\right)\leqq \operatorname{Var}\left(R\left(t+\frac{1}{m}\right)\right)\Big/\left(\mathsf{E}\left(R\left(t+\frac{1}{m}\right)\right)-\frac{1}{m}\right)^2\leqq$$

$$\leqq 14N^{-1/2}\left(\frac{1}{m}-7N^{-1/2}\right)^{-2}.$$

Since

$$\sqrt{N}\geqq m2N^{(1+\beta)/2(2+\beta)}\geqq 2mN^{\beta/(2+\beta)}\geqq 128m,$$

we have

$$\mathsf{P}\left(R\left(t+\frac{1}{m}\right)\leqq t\right)\leqq 14N^{-1/2}\left(\frac{128}{121}\right)^2 m^2\leqq 4N^{-(\beta/2(2+\beta))}.$$

By substituting all these into (8) we arrive at the upper estimation

$$\mathsf{P}(R_B\leqq t)-t^\beta\leqq 8N^{-\beta/(2+\beta)}.$$

Similarly, if $\frac{1}{m}<t<1$,

$$\mathsf{P}(R_B\leqq t)\geqq \mathsf{P}\left(S_B\leqq t-\frac{1}{m}\right)-\mathsf{P}\left(S_B\leqq t-\frac{1}{m},\ R_B>t\right)\geqq$$

$$\geqq \mathsf{P}\left(S_B\leqq t-\frac{1}{m}\right)-\mathsf{P}\left(R\left(t-\frac{1}{m}\right)>t\right)\geqq t^\beta-m^{-\beta}-2N^{-1/2}-4N^{-(\beta/2(2+\beta))}\geqq$$

$$\geqq t^\beta-8N^{-(\beta/2(2+\beta))}.$$

For $1-\frac{1}{m}\leqq t\leqq 1$ we can use the trivial upper estimate 1 as well as the lower estimate 0 in the case $0<t\leqq\frac{1}{m}$. The proof is complete.

5. Proof of Theorem 2

(a) The proof will follow the line of reasoning of [3], Section 5.

In the sequel x_n will always denote $d^n\log\frac{1}{t_n}$, that is, t_n is derived from x_n in the same way as S_n is from $T_n=T(B_n)$.

(i) Suppose $\sum_{n=1}^{\infty} t_n^{\beta_n}<\infty$. Since by (6)

$$\left|\sum_{n=1}^{\infty}\mathsf{P}(S_n\leqq t_n)-\sum_{n=1}^{\infty} t_n^{\beta_n}\right|\leqq 2\sum_{n=1}^{\infty} nd^{-n}<\infty \tag{9}$$

the Borel—Cantelli lemma implies that $S_n>t_n$ with only finitely many exceptions for n.

Suppose $\sum_{n=1}^{\infty} t_n^{\beta_n}=\infty$. We will use the Erdős—Rényi generalization of the Borel—Cantelli lemma. It asserts that if $\sum_{n=1}^{\infty} \mathsf{P}(A_n)=\infty$ and

$$\liminf_{n\to\infty} \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} \mathsf{P}(A_i\cap A_j)}{\left(\sum_{i=1}^{n} \mathsf{P}(A_i)\right)^2} \leqq 1 \tag{10}$$

holds for a sequence of events A_n, then, with probability 1, infinitely many of the events occur.

Now consider the events $A_n=\{S_n\leqq t_n\}=\{T_n\geqq x_n\}$. We can suppose that $\mathsf{P}(A_n)\to 0$, otherwise we could find a subsequence of integers n' along which $P(A_{n'})\to\to c>0$. Hence $\mathsf{P}(\limsup A_n)\geqq c$, and, since the limsup event is measurable with respect to all tail-σ-fields of the sequence of experiments, it must be of probability 1 by the 0—1 law.

Consider the double sum in the numerator of (10).

$$\sum_{i=1}^{n}\sum_{j=1}^{n} \mathsf{P}(A_i\cap A_j) = \sum_{i=1}^{n} \mathsf{P}(A_i) - 2\sum_{1\leqq i<j\leqq n} \mathsf{P}(A_i\cap A_j).$$

The first sum is negligible compared with the denominator. The second sum is to be divided into two parts according that $x_i\geqq x_j$ or $x_i<x_j$. Since by (3)

$$t^{\beta_i c_i} \leqq \mathsf{P}(S_i\geqq t) \leqq c_i\, t^{\beta_i}, \tag{11}$$

where $c_i=1+2id^{-i}\leqq 2$, in the first case we have

$$\mathsf{P}(A_i\cap A_j) \leqq \mathsf{P}(A_i) \leqq \mathsf{P}(T_i\geqq x_j) = \mathsf{P}\left(S_i\leqq t_j^{d^{i-j}}\right) \leqq 2t_j^{\beta_i d^{j-i}}.$$

Here $\beta_i d^{j-i}\geqq(d-1)(j-1)\geqq\beta_j(j-i)$, hence

$$2\sum_{\substack{1\leqq i<j\leqq n\\ x_i\geqq x_j}} \mathsf{P}(A_i\cap A_j) \leqq 4\sum_{j=2}^{n}\sum_{i=1}^{j-1} (t_j^{\beta_j})^{j-i} \leqq \sum_{j=2}^{n} t_j^{\beta_j}\frac{4}{1-t_j^{\beta_j}},$$

which is again negligible compared with the denominator of (10).

In the second case

$$\begin{aligned}\mathsf{P}(A_i\cap A_j) &= \mathsf{P}(T_i\geqq x_i \text{ and } T_j\geqq x_j) \leqq \mathsf{P}(T_i\geqq x_i)\,\mathsf{P}(T_j\geqq x_j-x_i) \leqq\\ &\leqq \frac{\mathsf{P}(T_i\geqq x_i)(\mathsf{P}(T_j\geqq x_j)+j\,d^{-j})}{\mathsf{P}(T_j\geqq x_i)}.\end{aligned} \tag{12}$$

On the one hand, by (12),

$$\mathsf{P}(A_i\cap A_j) \geqq \left(\mathsf{P}(A_i)P(A_j)+j\,d^{-j}\right) t_i^{-\beta_j c_j d^{i-j}} \leqq \left(\mathsf{P}(A_i)\mathsf{P}(A_j)+j\,d^{-j}\right) c_{ij}^2,$$

where $c_{ij}=t_i^{-d^{i-j}}$. On the other hand,

$$\begin{aligned}\mathsf{P}(A_i\cap A_j) &\leqq \mathsf{P}(A_i)\,\mathsf{P}(T_j\geqq x_j-x_i) \leqq c_i\,c_j\,t_i^{\beta_i}\exp\left(-\beta_j\,d^{-j}(x_j-x_i)\right) \leqq\\ &\leqq 4t_j^{\beta_j}\,t_i^{\beta_i-\beta_j d^{i-j}} \leqq 4t_j^{\beta_j}\,c_{ij}^{-(i-j-1)}.\end{aligned}$$

The contribution of terms for which $c_{ij}<1+\delta$ with a small positive δ is majorized by

$$(1+\delta)^2((\sum_{i=1}^{n} \mathsf{P}(A_i))^2+\sum_{j=1}^{n} j^2 d^{-j}).$$

The sum of terms for which $c_{ij}\geqq 1+\delta$ is relatively negligible, since

$$\sum_{j=2}^{n} t_j^{\beta_j} \sum_{i=1}^{j-1} c_{ij}^{-(j-i-1)} \leqq \frac{1}{\delta} \sum_{j=2}^{n} t_j^{\beta_j}.$$

Hence (10) follows.

(ii) Suppose $\sum_{n=1}^{\infty} u_n<\infty$. Let $t_n=1-u_n$ and $x_n=d^n \log \frac{1}{t_n}$. Since

$$|\mathsf{P}(S_n \geqq t_n)-(1-t_n^{\beta_n})| \leqq 2nd^{-n}; \tag{13}$$

further, for $u, \beta\in(0, 1)$

$$\beta u \leqq 1-t^{\beta} \leqq u \quad \text{and} \quad 1-t^{\beta} \leqq \beta \frac{u}{1-u} \qquad (t=1-u), \tag{14}$$

we have

$$\sum_{n=1}^{\infty} \mathsf{P}(S_n \geqq t_n) < \infty.$$

Thus the Borel—Cantelli lemma gives that $S_n<t_n$ with only finitely many exceptions.

Suppose conversely that $\sum_{n=1}^{\infty} u_n=\infty$. From (13)—(14) we can see that

$$\sum_{i=1}^{n} \mathsf{P}(S_i \geqq t_i) \asymp \sum_{i=1}^{n} u_i$$

(the ratio of the two sequences keeps bounded from 0 and ∞).

As in (i), we want to apply the Erdős—Rényi lemma to the events $A_n=\{S_n\geqq t_n\}=$ $=\{T_n\leqq x_n\}$.

Again, we can suppose $\mathsf{P}(A_n)\to 0$, which means $u_n\to 0$. Dealing with the sum $\sum_{1\leqq i<j\leqq n} \mathsf{P}(A_i\cap A_j)$ we treat separately the pairs (i,j) with $x_i\geqq x_j$ from those with $x_i<x_j$. In the first case

$$\mathsf{P}(A_i\cap A_j) \leqq \mathsf{P}(A_j) \leqq \mathsf{P}(T_j \leqq x_i),$$

while in the second case by (12)

$$\begin{aligned}\mathsf{P}(A_i\cap A_j) &= \mathsf{P}(A_i)\mathsf{P}(A_j)+\mathsf{P}(\bar{A}_i\cap \bar{A}_j)-\mathsf{P}(\bar{A}_i)\mathsf{P}(\bar{A}_j) \leqq \\ &\leqq \mathsf{P}(A_i)\mathsf{P}(A_j)+\mathsf{P}(T_i > x_i)\mathsf{P}(x_j-x_i < T_j \leqq x_j) \leqq \\ &\leqq \mathsf{P}(A_i)\mathsf{P}(A_j)+\mathsf{P}(T_j \leqq x_i).\end{aligned}$$

Here by (6) and (14)

$$\mathsf{P}(T_j \leqq x_i) = \mathsf{P}(S_j \geqq t_i^{d^{i-j}}) \leqq 1-t_i^{d^{i-j}}+2j\,d^{-j} \leqq d^{i-j} \frac{u_i}{1-u_i}+2j\,d^{-j},$$

hence

$$\sum_{1\le i<j\le n} \mathsf{P}(T_j \le x_i) \le \sum_{i=1}^{n} \frac{u_i}{1-u_i}\cdot\frac{1}{d-1} + \sum_{j=1}^{n} 2j^2 d^{-j} = O\Big(\sum_{i=1}^{n} u_i\Big),$$

which is negligible compared with $\big(\sum_{i=1}^{n} \mathsf{P}(A_i)\big)^2$. Now (10) immediately follows, completing the proof.

(b) Let $\varepsilon_n = N_n^{-1/6(\beta_n+1)}$. By Theorem 1(a) and the Markov inequality

$$\mathsf{P}(|R_n - S_n| \ge \varepsilon_n) \le 12 N_n^{-1/6}\, \varepsilon_n^{-1} = 12 N^{n-(\beta_n/6(\beta_n+1))}.$$

The sum of these probabilities converges by the assumption, hence the Borel—Cantelli lemma implies that with probability 1 $|R_n - S_n| < \varepsilon_n$ for every n large enough.

Next, let (t_n) and (t_n') be positive sequences for which $|t_n - t_n'| \le \varepsilon_n$. Then the series $\sum_n t_n^{\beta_n}$ and $\sum_n (t_n')^{\beta_n}$ are equiconvergent, since

$$\left|t_n^{\beta_n} - (t_n')^{\beta_n}\right| \le |t_n - t_n'|^{\beta_n} \le N^{-(\beta_n/6(1+\beta_n))}.$$

Clearly, the same is true for the series $\sum_n t_n$ and $\sum_n t_n'$.

Now the assertion of part (b) immediately follows. For instance, suppose that $\sum_n t_n^{\beta_n} < \infty$. Then $\sum_n (t_n + \varepsilon_n)^{\beta_n} < \infty$ as well, and by part (a), $S_n > t_n + \varepsilon_n$ for every n large enough, hence $R_n > t_n$ with only finitely many exceptions. All the other cases can be treated similarly.

REFERENCES

[1] ERDŐS, P., Personal communication.

[2] ERDŐS, P. and RÉVÉSZ, P., On the length of the longest head-run, *Topics in Information Theory* (Keszthely, Hungary, 1975), Colloq. Math. Soc. J. Bolyai, Vol. 16, North-Holland, Amsterdam, 1977, 219—228. *MR* 57 #17788.

[3] GUIBAS, L. J. and ODLYZKO, A. M., Long repetitive patterns in random sequences, *Z. Wahrscheinlichkeitstheorie Verw. Gebiete* **53** (1980), 241—262. *MR* **81m**: 60047.

[4] LI, S. R., A martingale approach to the study of occurence of sequence patterns in repeated experiments, *Ann. Probab.* **8** (1980), 1171—1176. *MR* **82c**: 60018

[5] MÓRI, T. F., Large deviation results for waiting times in repeated experiments, *Acta Math. Hungar.* **45** (1985), 213—221. *MR* **86c**: 60017.

[6] MÓRI, T. F., On the waiting time till each of some given patterns occurs as a run, *Probab. Th. Rel. Fields* (to appear).

[7] MÓRI, T. F., More on the waiting time till each of some given patterns occurs as a run (to appear).

[8] MÓRI, T. F., Maximum waiting time when the size of the alphabet increases, *Mathematical Statistics and Probability Theory* (Bad Tatzmannsdorf, Austria, 1986), Vol. B, Reidel, Dordrecht, 1987, 169—178.

(Received February 13 1987)

EÖTVÖS LORÁND TUDOMÁNYEGYETEM
TERMÉSZETTUDOMÁNYI KAR
VALÓSZÍNŰSÉGSZÁMÍTÁSI TANSZÉK
MÚZEUM KRT. 6—8
H—1088 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 365—375

ON AN OPERATIONAL DIFFERENTIAL EQUATION SYSTEM

TAMÁS FÉNYES

Introduction

In the paper [1] we have ds css ed the algebraic

$$(1) \qquad Dx_\mu - q \sum_{k=1}^{m} a_{\mu k} x_k = \gamma_\mu, \qquad \mu = 1, 2, \dots, m; \ m > 1$$

differential equation system defined in the Mikusiński operator field M based on the convolution product of the continuous (or locally integrable) functions. In (1) x_μ $(\mu=1, 2, \dots, m)$ is the unknown solution, if it exists, D is the symbol of the algebraic derivative, $a_{\mu k}$ are arbitrarily given real numbers, $q \in M$ is a given operator being the sum of an arbitrary number and of an arbitrary locally integrable function defined on $0 \leqq t < \infty$, the operators γ_μ are also arbitrarily given locally integrable functions defined on the interval $t \geqq 0$.

In the present paper we shall give an operational treatment of (1) in the discrete operator field M_D based on the Dirichlet product of functions defined on the set of the natural numbers. In the discrete operational calculus, the numbers can be algebraically identified with special discrete functions, so we assume that q, γ_μ are arbitrarily given real-valued functions defined on the set of the natural numbers. In the convolution ring based on the Dirichlet product (1) is equivalent to the following system of functional equations

$$(2) \qquad \log n \cdot x_\mu(n) + \sum_{k=1}^{m} a_{\mu k} \sum_{\nu/n} q(\nu)\, x_k\left(\frac{n}{\nu}\right) + \gamma_\mu(n) = 0; \qquad \mu = 1, 2, \dots, m.$$

We assume that the eigenvalues of the matrix $A=[a_{\mu k}]$ are all distinct.

For the theory and applications of the discrete operational calculus see [2], [3], [4]. In Chapter 1 we briefly summarize the results of the paper [2], giving some generalizations of them, Chapter 2 contains the operational theory of the differential equation system (1).

In what follows Z will denote the set of natural numbers.

Research (partially) supported by the National Foundation for Scientific Research Grant no. 6032/6319.

1980 *Mathematics Subject Classification* (1985 Revision). Primary 44A40.

Key words and phrases. Mikusiński's operator calculus, operational differential equation system.

§ 1. Discrete Mikusiński operators based on the Dirichlet product

Let $a=\{a(n)\}$ be an arbitrary real-valued function defined on Z. The symbol $a(n)$ denotes the value of this function for arbitrary fixed n.

Let E denote the set of the discrete functions. If we introduce in E the following two operations

(i) $a+b$: $\{a(n)\}+\{b(n)\} = \{a(n)+b(n)\}$, addition

(ii) $ab = \{a(n)\}\{b(n)\} = \left\{\sum_{\nu/n} a(\nu)\, b\left(\frac{n}{\nu}\right)\right\}$, multiplication,

then E becomes a commutative ring without divisor of zero and can be extended to a quotient field. This is called the discrete Mikusiński operator field and is denoted by M_D. The elements of M_D are called M_D-operators.

The definition and properties of the "discrete" Dirac function.

We define the discrete Dirac function by

$$\delta(N) = \{\delta(n, N)\},$$

where

$$\delta(n, N) = \begin{cases} 0 & \text{for } n \neq N \\ 1 & \text{for } n = N \end{cases} \qquad N \in Z.$$

For later purposes we enumerate some properties of the Dirac function.

PROPERTY 1. $\delta(N)\{a(n)\}=\{b(n)\}$,

$$b(n) = \begin{cases} a\left(\frac{n}{N}\right) & \text{for } N|n, \\ 0 & \text{otherwise.} \end{cases} \tag{1.1}$$

$$\delta(N_1)\,\delta(N_2) = \delta(N_1 N_2), \qquad N_1, N_2 \in Z. \tag{1.2}$$

PROPERTY 2.

$$x = \frac{\{a(n)\}}{\delta(N)} \in E, \qquad N \in Z \tag{1.3}$$

holds if and only if

$$a(n) = 0 \tag{1.4}$$

for those values of n for which N is not a divisor of n. If (1.4) holds, then

$$x = \{a(nN)\}. \tag{1.5}$$

The field K of the real or complex numbers can be embedded isomorphically into the operator field M_D. The common unit element of K, E, M_D is the function $\delta(1)$ and we write

$$\delta(1) = 1.$$

Moreover,

$$c\delta(1) = c, \quad c\{a(n)\} = \{ca(n)\};$$

for every $c \in K$, and every $a \in E$.

Every operator of the form

$$x = \frac{\{a(n)\}}{\{b(n)\}}$$

is a function if $b(1) \neq 0$.

The operator function $\delta(\varepsilon)$.

For arbitrary rational number $\varepsilon = \dfrac{N_1}{N_2}$ we define

$$\delta(\varepsilon) = \frac{\delta(N_1)}{\delta(N_2)}. \tag{1.6}$$

From this definition it follows that for $\varepsilon = N$ we have

$$\delta(\varepsilon) = \delta(N) = \{\delta(n, N)\}.$$

If

$$\frac{N_1}{N_2} = \frac{N_3}{N_4},$$

then

$$\delta\left(\frac{N_1}{N_2}\right) = \delta\left(\frac{N_3}{N_4}\right)$$

holds.

PROPERTY 3. Let α, β be arbitrary positive rational numbers, then

$$\delta(\alpha)\delta(\beta) = \delta(\alpha\beta) \tag{1.7}$$

and it is easily seen that

$$\delta\left(\frac{1}{\alpha}\right) = \frac{1}{\delta(\alpha)} \tag{1.8}$$

is also true.

The definition of the ring E^*

Let $E^* \subset M_D$ be the subset of M_D whose elements are of the form

$$x = \frac{a}{\delta(N)}, \qquad N \in Z, \quad a \in E. \tag{1.9}$$

E^* is a ring and, by choosing $N = 1$, we have

$$E \subset E^*.$$

PROPERTY 4. Obviously,

$$x = \frac{a}{\delta(\varepsilon)} \in E^*, \quad \varepsilon = \frac{N_1}{N_2} \quad (N_1, N_2 \text{ are relatively primes}).$$

Moreover, $x \in E$ if and only if

$$a(n) = 0$$

for those values of n for which N_1 is not a divisor of n. If the condition is satisfied, we have

$$x(n) = \begin{cases} a\left(\dfrac{n N_1}{N_2}\right) & \text{for} \quad N_2 | n, \\ 0 & \text{otherwise.} \end{cases}$$

Definition of the convergence in the ring E.

Let $a_k \in E$, $(k = 1, 2, \dots)$ be an infinite sequence of functions. By definition

(1.10) $$\lim_{k \to \infty} \{a_k(n)\} = \{a(n)\}$$

if for every fixed n

$$\lim_{k \to \infty} a_k(n) = a(n)$$

(see [5]). This convergence can be extended to infinite series of functions as usual. Let

$$f(z) = \sum_{k=0}^{\infty} \beta_k z^k, \quad \beta_k \in K$$

be an arbitrary entire function of the complex variable z. Then

(1.11) $$f(a) = \sum_{k=0}^{\infty} \beta_k \{a(n)\}^k, \quad a \in E, \quad a^0 = 1$$

holds in the sense of the convergence defined above. We have

$$e^a = \sum_{k=0}^{\infty} \frac{a^k}{k!}, \quad a \in E, \quad a^0 = 1$$

having the property

$$e^a e^b = e^{a+b}, \quad a, b \in E,$$

moreover, if we write

$$e^a = \{e_a(n)\},$$

so

(1.12) $$e_a(1) = e^{a(1)}$$

holds.

Moreover let

$$\sum_{k=0}^{\infty} \gamma_k z^k, \quad \gamma_k \in K$$

be an arbitrary *formal* infinite series and let $a \in E$ an arbitrary function with $a(1)=0$. Then

$$\sum_{k=0}^{\infty} \gamma_k a^k$$

also converges in the sense of convergence defined above.

The algebraic derivation and integration (see also [4]).

For the sake of easy reading we recapitulate some definitions and facts of the algebraic derivation and integration.

$$D(a) = \{-\log n \cdot a(n)\}, \quad a \in E \tag{1.13}$$

$$D\left(\frac{a}{b}\right) = \frac{b\,D(a) - a\,D(b)}{b^2}, \quad a, b \in E, \quad \frac{a}{b} \in M_D. \tag{1.13'}$$

PROPERTY 5.

$$D\left[\frac{a}{\delta(\varepsilon)}\right] = \frac{\left\{-\log \frac{n}{\varepsilon} \cdot a(n)\right\}}{\delta(\varepsilon)} \in E^*, \quad a \in E, \quad \varepsilon = \frac{N_1}{N_2}. \tag{1.14}$$

$$D[\delta(\varepsilon)] = -\log \varepsilon \cdot \delta(\varepsilon). \tag{1.14'}$$

PROPERTY 6.

$$D(e^a) = D(a)\,e^a, \quad a \in E. \tag{1.15}$$

If for a given $x \in M_D$ there exists a $y \in M_D$ such that

$$D(y) = x$$

we say that x is algebraic integrable and we write

$$y = \int x.$$

PROPERTY 7. If $x \in M_D$ and

$$D(x) = 0$$

then x is an arbitrary complex number.

Two algebraic integrals of an operator may differ only by an arbitrary number.

The algebraic differentiation and integration is a linear operation over the field of the real (complex) numbers.

PROPERTY 8. The operator

$$x = \frac{a}{\delta(\varepsilon)}, \quad a \in E, \quad \varepsilon = \frac{N_1}{N_2} \tag{1.16}$$

is algebraic integrable in M_D if and only if either $\varepsilon \neq N$, $N \in Z$, or $\varepsilon = N$ and $a(N)=0$ holds true. Every algebraic integral of (1.16) belonging to E^* is given by

$$\int \frac{a}{\delta(\varepsilon)} = \frac{\left\{-\dfrac{a(n)}{\log \dfrac{n}{\varepsilon}}\right\}}{\delta(\varepsilon)} + c, \quad c \in K \tag{1.17}$$

where in the case of $\varepsilon = N$ the symbol

$$\frac{a(N)}{\log \dfrac{N}{N}}$$

denotes an arbitrary real (complex) number. We shall choose this to be null.

For $\varepsilon = 1$ we have that a is integrable if and only if $a(1)=0$, and

$$\int a = \left\{-\frac{a(n)}{\log n}\right\} + c, \quad c \in K.$$

Let us consider the differential equation

$$D(x) - fx = h, \quad f \in E, h \in E \tag{1.18}$$

with respect to which the following theorem holds.

THEOREM 1 (see [2], [3]). *The homogeneous equation*

$$D(x) - fx = 0 \tag{1.19}$$

has a nontrivial solution in M_D if and only if $f(1)$ is real and $\alpha = e^{-f(1)}$ is rational.

The general solution of (1.19) is of the form

$$x = c\delta(\alpha) \exp\left[\int (f - f(1))\right], \quad c \in K \tag{1.20}$$

being a function for $c \neq 0$ if and only if α is natural. (1.18) has a solution $x_p \in M_D$ if and only if one of the following conditions holds:

(i) $f(1)$ is not real or

$$\text{(ii) } f(1) \text{ is real and } \alpha = e^{-f(1)} \notin Z \tag{1.20'}$$

or

$$\text{(iii) } f(1) \text{ is real}, \alpha = e^{-f(1)} \in Z, \quad H(\alpha) = 0,$$

where

$$H = \{H(n)\} = \{h(n)\} e^{-\int (f - f(1))}.$$

Moreover

$$x_p = \left\{\frac{-H(n)}{\log n + f(1)}\right\} \exp\left[\int (f - f(1))\right] \in E, \tag{1.21}$$

where in the case of (iii) the symbol

$$\frac{H(\alpha)}{\log\alpha+f(1)}=\frac{0}{0} \tag{1.22}$$

denotes the number zero.

§ 2. On the differential equation system (1)

Let us consider the differential equation system

$$D(x_\mu)-q\sum_{k=1}^{m}a_{\mu k}x_k=\gamma_\mu,\quad \mu=1,2,\dots,m;\ m>1. \tag{2.1}$$

We say that (2.1) has an operational solution if there exist M_D operators x_μ satisfying the system (2.1). If x_μ is a discrete function for every value of μ, we say that (2.1) has a function solution.

Let us denote by λ_j $(j=1,2,\dots,m)$ the eigenvalues of the matrix A, and the eigenvector of the transpose corresponding to λ_j, by

$$\omega_j=\begin{bmatrix}\omega_{j1}\\ \omega_{j2}\\ \vdots\\ \omega_{jm}\end{bmatrix},\quad j=1,2,\dots,m.$$

Multiplying the first equation of (2.1) by ω_{j1} the second equation by ω_{j2}, and so on, the m-th equation by ω_{jm}, and adding the equations so obtained we have

$$\sum_{\mu=1}^{m}\omega_{j\mu}Dx_\mu-q\sum_{\mu=1}^{m}\omega_{j\mu}\sum_{k=1}^{m}a_{\mu k}x_k=\sum_{\mu=1}^{m}\omega_{j\mu}\gamma_\mu \tag{2.2}$$

and

$$\sum_{\mu=1}^{m}\omega_{j\mu}Dx_\mu-q\sum_{k=1}^{m}x_k\sum_{\mu=1}^{m}a_{\mu k}\omega_{j\mu}=\sum_{\mu=1}^{m}\omega_{j\mu}\gamma_\mu. \tag{2.3}$$

Since

$$\sum_{\mu=1}^{m}a_{\mu k}\omega_{j\mu}=\lambda_j\omega_{jk}$$

so

$$\sum_{\mu=1}^{m}\omega_{j\mu}Dx_\mu-q\lambda_j\sum_{k=1}^{m}\omega_{jk}x_k=\sum_{\mu=1}^{m}\omega_{j\mu}\gamma_\mu. \tag{2.4}$$

Introducing the operators

$$z_j=\sum_{\mu=1}^{m}\omega_{j\mu}x_\mu,\qquad f_j=\sum_{\mu=1}^{m}\omega_{j\mu}\gamma_\mu$$

we get

$$Dz_j-q\lambda_j z_j=f_j,\qquad j=1,2,\dots,m. \tag{2.5}$$

In this way the system (2.1) has been reduced to the equations (2.5). If we know the solutions $z_j \in M_D$ $(j=1, 2, \dots, m)$ of (2.5), then by solving the algebraic equation system

$$\sum_{\mu=1}^{m} \omega_{j\mu} x_\mu = z_j, \qquad j = 1, 2, \dots, m$$

we obtain an operational solution of (2.1) of the form

$$x_\mu = \sum_{j=1}^{m} \beta_{j\mu} z_j, \qquad \mu = 1, 2, \dots, m; \ \beta_{j\mu} \in K. \tag{2.6}$$

So we get the following

Lemma 1. (2.1) *has an operational solution if and only if each differential equation of the system* (2.5) *has an operational solution.*

Proof. Indeed, if (2.1) has an operational solution, then the operators

$$z_j = \sum_{\mu=1}^{m} \omega_{j\mu} x_\mu$$

satisfy the differential equations (2.5). Conversely, if the operators z_j satisfy (2.5), then (2.6) is an operational solution of (2.1).

A similar statement holds for the function solution of (2.1).

Lemma 2. (2.1) *has a function solution if and only if each differential equation of the system* (2.5) *has a function solution.*

In the following we shall deal with the homogeneous system

$$Dx_\mu - q \sum_{k=1}^{m} a_{\mu k} x_k = 0, \qquad \mu = 1, 2, \dots, m. \tag{2.7}$$

By (2.5) we have

$$Dz_j - q\lambda_j z_j = 0, \quad j = 1, 2, \dots, m. \tag{2.8}$$

Let Q denote the set of those indices j for which every differential equation

$$Dz_j - q\lambda_j z_j = 0, \quad j \in Q$$

has a nontrivial operational solution in M_D. The general solution is given by (1.20)

$$z_j = c_j \delta(\alpha_j) \exp\left[\int (\lambda_j q - \lambda_j q(1))\right], \qquad \alpha_j = e^{-\lambda_j q(1)}, \quad j \in Q, c_j \in K.$$

Taking into account Lemma 1 and (2.6) we obtain the general operational solution of (2.1) in the form of

$$x_\mu = \sum_{j \in Q} c_j \beta_{j\mu} \delta(\alpha_j) \exp\left[\int (\lambda_j q - \lambda_j q(1))\right], \quad \mu = 1, 2, \dots, m. \tag{2.9}$$

Let $Q^* \subset Q$ denote the set of those indices j for which every differential equation

$$Dz_j - q\lambda_j z_j = 0, \quad j \in Q^*$$

has a nontrivial solution in the ring E. By Lemma 2 and (2.6) we obtain the general function solution of (2.1) in the form

$$x_\mu = \sum_{j \in Q^*} c_j \beta_{j\mu} \delta(\alpha_j) \exp\left[\int (\lambda_j q - \lambda_j q(1))\right], \qquad \mu = 1, 2, \ldots, m. \tag{2.10}$$

So the following assertion holds true:

THEOREM 2. *Let the eigenvalues of the matrix $A = [a_{\mu k}]$ be distinct. The homogeneous system* (2.7) *has a nontrivial solution in M_D if and only if there exists an index j $(1 \leqq j \leqq m)$ for which $\lambda_j q(1)$ is real and $e^{-\lambda_j q(1)}$ is rational. Moreover,* (2.7) *has a nontrivial solution in the ring E if and only if there exists an index j for which $\lambda_j q(1)$ is real and $e^{-\lambda_j q(1)}$ is natural. The general solutions are given by* (2.9), (2.10).

REMARK. If the matrix A has complex eigenvalues, then the general function solution given by (2.10), is complex-valued in general. However, the general real-valued solution in E can be given by using the following procedure.

Let $Q^{**} \subset Q^*$ denote the set of those indices j for which

$$\lambda_j = \tau_j + i\sigma_j \qquad (\sigma_j \neq 0).$$

Let $q(1) = 0$ and so

$$\alpha_j = 1$$

for every j. By (2.10) we have

$$x_\mu = \sum_{j \in Q^*} c_j \beta_{j\mu} e^{\int \lambda_j q} = \sum_{j \in Q^{**}} c_j \beta_{j\mu} e^{\int \lambda_j q} + \sum_{j \in Q^* - Q^{**}} c_j \beta_{j\mu} e^{\int \lambda_j q}. \tag{2.11}$$

Let $j_1, j_2 \in Q^{**}$ such that

$$\lambda_{j_1} = \tau + i\sigma, \qquad \lambda_{j_2} = \tau - i\sigma.$$

It can easily be seen that the numbers $\beta_{j_1\mu}, \beta_{j_2\mu}$ are conjugate complex ones

$$\beta_{j_1\mu} = \varrho_\mu + i\varepsilon_\mu,$$

$$\beta_{j_2\mu} = \varrho_\mu - i\varepsilon_\mu, \qquad \mu = 1, 2, \ldots, m,$$

moreover, the exponential functions

$$e^{\int \lambda_{j_1} q}, \; e^{\int \lambda_{j_2} q}$$

are conjugate complex valued for every $n = 1, 2, \ldots$. Let us choose the coefficients c_{j_1}, c_{j_2} to be real, and let

$$c_{j_1} = c_{j_2} = \frac{c}{2}.$$

A particular real-valued solution of (2.7) is

$$x_{p\mu} = \frac{c}{2}(\varrho_\mu + i\varepsilon_\mu)\exp\int(\tau+i\sigma)\,q + \frac{c}{2}(\varrho_\mu - i\varepsilon_\mu)\exp\int(\tau-i\sigma)q =$$

$$= \frac{c}{2}(\varrho_\mu + i\varepsilon_\mu)\exp\int\tau q\left(\cos\int\sigma q + i\sin\int\sigma q\right) +$$

$$+\frac{c}{2}(\varrho_\mu - i\varepsilon_\mu)\exp\int\tau q\left(\cos\int\sigma q - i\sin\int\sigma q\right) =$$

$$= c\exp\int\tau q\left(\varrho_\mu\cos\int\sigma q - \varepsilon_\mu\sin\int\sigma q\right), \quad \mu = 1, 2, \dots, m.$$

Another linearly independent particular real-valued solution of (2.7) can be obtained by choosing

$$c_{j_1} = \frac{D}{2i}, \quad c_{j_2} = -\frac{D}{2i}, \quad D \text{ real}.$$

$$x_{p\mu} = \frac{D}{2i}(\varrho_\mu + i\varepsilon_\mu)\exp\int(\tau+i\sigma)\,q - \frac{D}{2i}(\varrho_\mu - i\varepsilon_\mu)\exp\int(\tau-i\sigma)\,q =$$

$$= \frac{D}{2i}(\varrho_\mu + i\varepsilon_\mu)e^{\int\tau q}\left(\cos\int\sigma q + i\sin\int\sigma q\right) -$$

$$-\frac{D}{2i}(\varrho_\mu - i\varepsilon_\mu)e^{\int\tau q}\left(\cos\int\sigma q - i\sin\int\sigma q\right) =$$

$$= De^{\int\tau q}\left(\varrho_\mu\sin\int\sigma q + \varepsilon_\mu\cos\int\sigma q\right), \quad \mu = 1, 2, \dots, m.$$

Performing this procedure for every pair of indices $j \in Q^{**}$ for which the corresponding eigenvalues of A are conjugate complex numbers and by choosing those coefficients c_j to be real for which $j \in Q^* - Q^{**}$, we get the general real-valued solution of (2.7).

Let us now discuss the inhomogeneous system (2.1). As we have seen in the foregoing (2.1) can be reduced to

$$Dz_j - q\lambda_j z_j = f_j, \quad j = 1, 2, \dots, m. \tag{2.12}$$

The existence criteria of the solutions of (2.12) and the explicit solution formulae are given in Chapter 1 (see (1.20′)). If for a fixed j there exists an operational solution of (2.12), then a particular solution z_{pj} is given by (1.18), (1.21), (1.22) in the form

$$z_{pj} = \left\{-\frac{H_j(n)}{\log n + \lambda_j q(1)}\right\}\exp\left[\int(\lambda_j q - \lambda_j q(1))\right] \in E, \tag{2.13}$$

where

$$H_j = f_j\exp\left[-\int(\lambda_j q - \lambda_j q(1))\right].$$

Taking into account Lemma 2 and (2.6), we obtain a particular function solution of the inhomogeneous system (2.1) of the form

$$x_{p\mu} = \sum_{j=1}^{m} \beta_{j\mu} z_{pj} \in E, \quad \mu = 1, 2, \ldots, m. \tag{2.14}$$

It can be easily shown that (2.14) is real-valued. The above discussions show that for the function solutions of (2.1) the following theorem holds:

ALTERNATIVE THEOREM. *Let the eigenvalues of the matrix A be distinct. If the homogeneous system* (2.7) *has only the trivial zero solution in E, then the inhomogeneous system* (2.1) *has exactly one solution in E. If the homogeneous system has nontrivial solutions in E, then the inhomogeneous system either has no solution in E, or has infinitely many solutions in E, according to the functions γ_μ.*

The general operational solution of (2.1) can be written as the sum of the general operational solution of the corresponding homogeneous system (2.7) and of a particular solution of (2.1). Every solution belongs to the ring E^*, more precisely, every element of any solution vector belongs to E^*.

REFERENCES

[1] FÉNYES, T. and KOSIK, P., Az operátortest egy algebrai differenciálegyenlet-rendszeréről, *Mat. Lapok* **30** (1978/82), 155—166. *MR* **85j**: 44008.

[2] FÉNYES, T. and SZILÁRD, K., Über diskrete Mikusinskische Operatoren, die auf Grund der Dirichletschen Produktenformel erzeugt werden, *Studia Sci. Math. Hungar.* **11** (1976), 181—199. *MR* **81b**: 44014b.

[3] FÉNYES, T., On a discrete nonlinear operational differential equation system based on the Dirichlet product, *Studia Sci. Math. Hungar.* **22** (1987), 471—484.

[4] GESZTELYI, E., The application of the operational calculus in the theory of numbers, *Number Theory* (Colloq., Math. Soc. János Bolyai, Debrecen, 1968), Colloq. Math. Soc. J. Bolyai, Vol. **2**, North-Holland, Amsterdam, 1970, 51—104. *MR* **42** # 5922.

[5] BUTZER, P. and SCHULTE, H., *Ein Operatorenkalkül zur Lösung gewöhnlicher und partieller Differenzengleichungssysteme von Funktionen diskreter Veränderlicher und seine Anwendungen*, Forschungsberichte des Landes Nordrhein—Westfalen, Nr. 1557, Westdeutscher Verlag, Köln und Opladen, 1965. *MR* **33** # 511.

(Received June 8, 1987)

MTA MATEMATIKAI KUTATÓ INTÉZETE
POSTAFIÓK 127
H—1364 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 377—383

VOLLSTÄNDIG ZIRKULÄRE KURVEN *n*-TER ORDNUNG DER ISOTROPEN EBENE

HANS SACHS

Meinem geschätzten Freund und Kollegen, Herrn O. Prof. Dr. J. Wills zum 50. Geburtstag gewidmet

In [2] entdeckte D. Palman eine interessante *Potenzeigenschaft* der *Tridens-Kurve*

$$xy = ax^3 + bx^2 + cx + d \tag{1}$$

in der isotropen Ebene $\mathbf{I}_2$. Unter einer *isotropen Ebene* versteht man hierbei eine projektive Ebene $\mathbf{P}_2(\mathbf{R})$, die über eine Absolutfigur $\{f, \mathbf{F}\}$ — bestehend aus einer Geraden f und einem mit f inzidenten Punkt $\mathbf{F}$ — metrisiert wird. Die interessante Geometrie dieser Ebene, welche erstmals von K. Strubecker in [4]—[7] entwickelt wurde, hat der Autor in der Monographie [3] systematisch dargestellt. Ziel dieser Note ist es, im ersten Abschnitt den *Potenzbegriff* von D. Palman auf algebraische Kurven n-ter Ordnung zu verallgemeinern und damit eine spezielle Kurvenklasse, die sogenannten *vollständig zirkulären Kurven* zu kennzeichnen; im zweiten Abschnitt der Abhandlung werden Eigenschaften dieser Kurven betrachtet und eine *geometrische Erzeugungsweise* einer speziellen Unterklasse angegeben. Wir benützen hierbei i.f. rein algebraische Methoden, wie sie auch vom *Jubilar* wiederholt in eleganten Abhandlungen eigesetzt wurden (vgl. z. B. [8]). Alle im folgenden verwendeten Begriffe der isotropen Geometrie können in [3] nachgelesen werden.

1. Wir bezeichnen mit $\{x, y\}$ affine Koordinaten in der isotropen Ebene $\mathbf{I}_2$ und mit $(x_0 : x_1 : x_2)$ die zugehörigen projektiven Koordinaten. Wird dann die *absolute Gerade* f durch $x_0 = 0$ erfaßt, und der *absolute Punkt* $\mathbf{F}$ durch $\mathbf{F}(0:0:1)$ beschrieben, dann legen wir den folgenden Betrachtungen die dreiparametrige isotrope Bewegungsgruppe $\mathfrak{B}_3$ als *Fundamentalgruppe* zu Grunde, welche koordinatenmäßig die Darstellung

$$\begin{aligned} \bar{x} &= a_1 + x \\ \bar{y} &= a_2 + a_3 x + y \end{aligned} \tag{2}$$

besitzt (vgl. [4,302]). Algebraische Kurven n-ter Ordnung $k^{(n)}$ beschreiben wir i.f. entweder durch

$$\sum_{j,k=0}^{j+k\leq n} a_{jk} x^j y^k = 0 \qquad \text{bzw. durch} \tag{3}$$

$$\sum_{i,j,k=0}^{i+j+k=n} a_{ijk} x_0^i x_1^j x_2^k = 0. \tag{4}$$

1980 *Mathematics Subject Classification* (1985 Revision). Primary 53A35; Secondary 51N35.
Key words and phrases. Isotropic plane, algebraic curves of degree n, total circulary, generalized Tridens curves.

Sei $k^{(n)}$ eine algebraische Kurve n-ter Ordnung in der komplex erweiterten isotropen Ebene $\mathbf{I}_2(\mathbf{C})$, $\mathbf{P}$ ein Punkt und g eine nichtisotrope Gerade, die mit $\mathbf{C}$ keinen Fernpunkt gemeinsam hat. Dann besitzen g und $k^{(n)}$ genau n Schnittpunkte $\{\mathbf{X}_1, \ldots, \mathbf{X}_n\}$ im algebraischen Sinn.

DEFINITION 1. Unter der *Potenz* des Punktes $\mathbf{P}$ in der Geraden g bezüglich der algebraischen Kurve $k^{(n)}$ versteht man das Produkt

$$\mathfrak{k}(P,g) = \prod_{i=1}^{n} d(P, X_i), \tag{5}$$

wobei $d(\mathbf{P}, \mathbf{X}_i)$ $(i=1, \ldots, n)$ die isotropen Abstände der Punkte $\mathbf{X}_i$ von $\mathbf{P}$ bezeichnen.

Wir beweisen zunächst den

SATZ 1. *Alle algebraischen Kurven n-ter Ordnung $k^{(n)}$ der isotropen Ebene $\mathbf{I}_2$, für welche die Potenz $\mathfrak{k}$ eines Punktes P nicht von der Lage der Geraden g durch P abhängt, besitzen die Gestalt*

$$\sum_{j,k=0}^{j+k\leqq n} a_{jk}\, x^j y^k = 0 \quad \textit{mit} \quad a_{n-\varrho\varrho} = 0 \quad \textit{für} \quad 1 \leqq \varrho \leqq n. \tag{6}$$

Für die Potenz $\mathfrak{k}(\mathbf{P})$ im Punkt $P(p_1, p_2)$ gilt

$$\mathfrak{k}(\mathbf{P}) = \frac{(-1)^n}{a_{no}} \sum_{j,k}^{j+k\leqq n} a_{jk}\, p_1^j p_2^k \quad \textit{mit} \quad a_{n-\varrho\varrho} = 0 \quad \textit{für} \quad 1 \leqq \varrho \leqq n. \tag{7}$$

BEWEIS. Da wir die Kurve $k^{(n)}$ als algebraische Kurve n-ter Ordnung in $\mathbf{I}_2$ voraussetzen, ist die Ferngerade f nicht Bestandteil von $k^{(n)}$. Hiermit existieren Geraden g durch $\mathbf{P}$, welche $k^{(n)}$ in eigentlichen Punkten $\{\mathbf{X}_1, \ldots, \mathbf{X}_n\}$ schneiden. Wird g in der Form $\{x=p_1+tv_1,\ y=p_2+tv_2\}$ mit $v_1 \neq 0$ parametrisiert, so werden die Schnittpunkte $\{\mathbf{X}_1, \ldots, \mathbf{X}_n\}$ durch die Nullstellen $\{t_1, \ldots, t_n\}$ des Polynoms

$$\sum_{j,k=0}^{j+k\leqq n} a_{jk}(p_1+tv_1)^j (p_2+tv_2)^k = 0 \tag{8}$$

erfaßt. Nun gilt $\mathfrak{k}(\mathbf{P}, g) = t_1 \cdot t_2 \cdot \ldots \cdot t_n v_1^n$, und andererseits folgt aus (8) nach Vietà

$$t_1 \ldots t_n = (-1)^n \frac{\sum\limits_{j,k=0}^{j+k\leqq n} a_{jk}\, p_1^j p_2^k}{\sum\limits_{j+k=n} v_1^j v_2^k}. \tag{9}$$

Wird noch $\lambda := \dfrac{v_2}{v_1}$ gesetzt, so darf schließlich

$$\mathfrak{k}(\mathbf{P}) = (-1)^n \frac{\sum\limits_{j,k=0}^{j+k\leqq n} a_{jk}\, p_1^j p_2^k}{\sum\limits_{j+k=n} a_{jk} \lambda^k} \tag{10}$$

nicht von λ abhängen, was die Bedingung $a_{n-11} = a_{n-22} = \ldots = a_{0n} = 0$ nach sich zieht;

es gilt hingegen $a_{n0} \neq 0$. Diese Bedingung ist auch hinreichend und die Potenz $\mathfrak{k}(\mathbf{P})$ ist — wie man aus (10) folgert — durch (7) gegeben

w.z.z.w.

Rechnet man (6) auf projektive Koordinaten um und bringt man die Kurve $k^{(n)}$ mit der absoluten Geraden $f(x_0=0)$ zum Schnitt, so stellt sich $\mathbf{F}(0:0:1)$ als einziger Schnittpunkt ein, der n-fach zu zählen ist. In Analogie zur euklidischen Situation liegt nahe die

Definition 2. Eine algebraische Kurve $k^{(n)}$ n-ter Ordnung der isotropen Ebene $\mathbf{I} \subset \mathbf{P}_2$, welche mit der absoluten Geraden f in $\mathbf{F}$ einen genau n-fachen Schnittpunkt hat, heißt eine *vollständig zirkuläre Kurve* n-ter Ordnung der isotropen Ebene.

Nach Satz 1 sind die *vollständig zirkulären Kurven* n-ter Ordnung somit dadurch *gekennzeichnet,* daß die *Potenz* $\mathfrak{k}(\mathbf{P})$ eines Punktes $\mathbf{P}$ *nicht* von der Geraden g durch $\mathbf{P}$ *abhängt.*

Folgerungen. 1) Gemäß der Nebenbedingung in (6) dürfen bei einer vollständig zirkulären Kurve n-ter Ordnung folgende Glieder nicht in der Kurvengleichung auftreten

$$a_{n-11}\, x^{n-1}\, y,\ a_{n-22}\, x^{n-1}\, y^2,\ \ldots,\ a_{1n-1}\, xy^{n-1},\ a_{0n}\ y^n. \tag{11}$$

2) Wir betrachten die Fälle $1 \leqq n \leqq 3$: Für $n=1$ erhält man $a_{00}+a_{10}x=0$, d.h. *isotrope Geraden.* Für diese ist die Potenzeigenschaft unmittelbar einsichtig. Für $n=2$ findet man die Kurven

$$a_{00}+a_{10}\,x+a_{01}\,y+a_{20}\,x^2 = 0. \tag{12}$$

Für $a_{01} \neq 0$ liegen *isotrope Kreise* vor, für welche die Potenzeigenschaft schon lange bekannt ist (vgl. [4,348f]). Für $a_{01}=0$ treten reduzible Fälle auf, nämlich 2 isotrope Geraden, die reell oder konjugiert-komplex sein können oder eine isotrope Doppelgerade. Für $n=3$ bekommt man die Lösungskurven

$$a_{00}+a_{10}\,x+a_{01}\,y+a_{20}\,x^2+a_{11}\,xy+a_{02}\,y^2+a_{30}\,x^3 = 0. \tag{13}$$

Diese Kurven wurden von D. Palman in [2] ausführlich studiert. Abgesehen von reduziblen Fällen, zerfallen sie in 3 Hauptklassen, nämlich die *divergenten Parabeln*

$$y^2 = ax^3+bx^2+cx+d, \tag{14}$$

bei denen $\mathbf{F}$ ein Wendepunkt mit f als Wendetangente ist, die *kubischen Parabeln*

$$y = ax^3+bx+cx+d \tag{15}$$

mit $\mathbf{F}$ als Spitze, wobei f die Spitzentangente ist und die *Tridens-Kurven* (1), bei denen $\mathbf{F}$ ein Knoten mit f als einer Tangente ist. Die von D. Palman entdeckte Potenzeigenschaft der Tridens-Kurven kommt somit auch den divergenten und kubischen Parabeln zu, wie unsere allgemeinen Überlegungen zeigen.[1]

[1] Die vollständig zirkulären Kurven 4. Ordnung wurden inzwischen von D. Palman in [2a] systematisch untersucht.

2. Zur Herleitung einiger Eigenschaften von vollständig zirkulären Kurven geben wir zunächst die

DEFINITION 3. Sind $k_1^{(n)}$ und $k_2^{(m)}$ zwei vollständig zirkuläre Kurven der Ordnung n bzw. m in $\mathbf{I}_2$, dann versteht man unter der *Potenzkurve* $p(k_1^{(n)}, k_2^{(m)})$ *der beiden Kurven* $k_1^{(n)}$ und $k_2^{(m)}$ die Menge der Punkte $\mathbf{P}\in\mathbf{I}_2$, die bezüglich $k_1^{(n)}$ und $k_2^{(m)}$ gleiche Potenz besitzen. Die *Potenzkurve* $p(k_1^{(n)})$ *einer* vollständig zirkulären *Kurve* n-ter Ordnung in $\mathbf{I}_2$ ist die Menge der Punkte $\mathbf{P}$, die bezüglich $k_1^{(n)}$ konstante Potenz besitzen.

SATZ 2. *Sei* $n>m$ *und seien* $k_1^{(n)}$ *und* $k_2^{(m)}$ *zwei vollständig zirkuläre Kurven, dann ist die Potenzkurve* $p(k_1^{(n)}, k_2^{(m)})$ *eine vollständig zirkuläre Kurve n-ter Ordnung. Für* $n=m$ *ist* $p(k_1^{(n)}, k_2^{(m)})$ *eine Kurve maximal* $(n-1)$*-ter Ordnung, welche i.a. nicht vollständig zirkulär ist. Die Potenzkurve einer vollständig zirkulären Kurve* $k^{(n)}$ *ist wieder eine vollständig zirkuläre Kurve n-ter Ordnung.*

BEWEIS. Hat $\mathbf{P}$ die Koordinaten $\mathbf{P}(x, y)$, dann wird nach (7) die Potenzkurve $p(k_1^{(n)}, k_2^{(m)})$ durch $\mathfrak{k}_1(\mathbf{P})=\mathfrak{k}_2(\mathbf{P})$, d. h.

$$(16) \qquad \frac{(-1)^n}{a_{n0}} \sum_{j,k}^{j+k\leq n} a_{jk}x^j y^k = \frac{(-1)^m}{b_{m0}} \sum_{j,k}^{j+k\leq m} b_{jk}x^j y^k$$

mit $a_{n-\varrho\varrho}=0$ für $1\leq\varrho\leq n$ und $b_{m-\varrho\varrho}=0$ für $1\leq\varrho\leq m$ festgelegt, wobei die a_{jk} bzw. b_{jk} die Koeffizienten in den Kurvendarstellungen von $k_1^{(n)}$ bzw. $k_2^{(m)}$ bezeichnen. (16) ist für $n>m$ eine algebraische Kurve n-ter Ordnung, welche vollständig zirkulär ist, da die Glieder $a_{n-11}x^{n-1}y, \ldots, a_{0n}y^n$ in (16) nicht vorkommen.

Für $n=m$ fällt in (16) das Glied x^n heraus, sodaß eine Kurve von maximal $(n-1)$-ter Ordnung vorliegt. Diese muß nicht vollständig zirkulär sein, wie das bekannte Beispiel (vgl. [4, 349]) zweier *isotroper Kreise* $y=\mathbf{R}_i x^2+\alpha_i x+\beta_i$ $(i=1, 2)$ zeigt. Sind die beiden Kreise nicht kongruent $(\mathbf{R}_1\neq\mathbf{R}_2)$, dann ist p eine *nichtisotrope Potenzgerade,* sind hingegen die beiden Kreise kongruent und konzentrisch $(\mathbf{R}_1=\mathbf{R}_2, \alpha_1=\alpha_2)$, dann ist p eine isotrope Gerade. Die letzte Aussage folgt sofort aus (7) w.z.z.w.

Dieser und auch der folgende Satz verallgemeinert je eine Resultat aus [2].

SATZ 3. *Sind* $k_1^{(n)}$ *und* $k_2^{(m)}$ *zwei vollständig zirkuläre Kurven der Ordnungen n bzw. m, dann liegen alle Punkte P der isotropen Ebene, für die das Produkt der Potenzen bezüglich* $k_1^{(n)}$ *und* $k_2^{(m)}$ *konstant ist, auf einer vollständig zirkulären Kurve der Ordnung* $(n+m)$.

BEWEIS. Nach (7) gilt für $\mathbf{P}(x, y)$ wegen $\mathfrak{k}_1(\mathbf{P})\cdot\mathfrak{k}_2(\mathbf{P})=c_0=\text{konst.}$

$$(17) \qquad c_0 = \frac{(-1)^{n+m}}{a_{n0}\, m_{n0}} \sum_{j,k=0}^{j+k\leq n} a_{jk}x^j y^k \cdot \sum_{j,k=0}^{j+k\leq n} b_{jk}x^j y^k$$

mit $a_{n-\varrho\varrho}=0$ für $1\leq\varrho\leq n$ und $b_{m-\varrho\varrho}=0$ für $1\leq\varrho\leq m$, wenn die a_{jk} bzw. b_{jk} die Koeffizienten der Kurven $k_1^{(n)}$ bzw. $k_2^{(m)}$ bezeichnen. (17) ist eine algebraische Kurve der Ordnung $(n+m)$. Diese ist vollständig zirkulär, da in der Produktbildung (17) die Glieder $d_{n+m-11}x^{n+m-1}y, \ldots, d_{0n+m}y^{n+m}$ ersichtlich nicht vorkommen w.z.z.w.

Zur Klasse der vollständig zirkulären Kurven n-ter Ordnung gehören auch die *reinen Parabeln* $k^{(n)}$

$$y = Ax^n. \tag{18}$$

Wir geben eine *geometrische Deutung* des Koeffizienten $\mathbf{A}$ im Rahmen der Bewegungsgruppe $\mathfrak{B}_3$ im

SATZ 4. *Ist P ein Punkt, der nicht auf einer reinen Parabel $k^{(n)}$ (18) liegt und bezeichnet $\tilde{P}$ den Schnittpunkt der isotropen Geraden durch P mit $k^{(n)}$, dann gilt*

$$A = (-1)^{n+1}\frac{s(P, \tilde{P})}{\mathfrak{k}(P)}, \tag{19}$$

wobei $s(P, \tilde{P})$ die Spanne der beiden Punkte P und $\tilde{P}$ bezeichnet.

BEWEIS. Nach (7) gilt für $\mathbf{P}(x, y)$ die Beziehung $\mathbf{A}\mathfrak{k}(\mathbf{P})=(-1)^n[\mathbf{A}x^n-y]$ und andererseits findet man $\bar{\mathbf{P}}(x, \mathbf{A}x^n)$, d. h. $s(\bar{\mathbf{P}}, \mathbf{P})=y-\mathbf{A}x^n$, woraus (19) folgt

w.z.z.w.

Um den Begriff der Tridens-Kurve (1) zu verallgemeinern geben wir die

DEFINITION 4. Unter einer verallgemeinerten *Tridens-Kurve* $k^{(n)}$ n-ter Ordnung der isotropen Ebene versteht man ein Monoid, welches F als $(n-1)$-fachen Punkt besitzt, wobei die absolute Gerade f genau Tangente eines Zweiges von $k^{(n)}$ ist.

Nach einem bekannten Satz der algebraischen Geometrie (vgl. [1,74]) lassen sich verallgemeinerte Tridens-Kurven $k^{(n)}$ somit in der Gestalt

$$\alpha_{n-2}(x)\, y+\beta_n(x) = 0 \tag{20}$$

schreiben, wobei $\alpha_{n-2}(x)=a_{n-2}x^{n-2}+\ldots+a_0$ ein Polynom vom Grad $(n-2)$ und $\beta_n(x)=b_n x^n+\ldots+b_0$ ein Polynom n-ten Grades bezeichnet. Wir werden im folgenden stets den Koeffizienten a_{n-2} zu 1 normieren und sprechen dann von einem *normierten Tridens n-ter Ordnung*. Als Tridens 2. Ordnung stellt sich ein isotroper Kreis ein. Für $n=3$ erhält man — nach Anwendung einer Schiebung in x-Richtung — genau (1), sodaß (20) als Verallgemeinerung von (1) angesehen werden kann. Die Nullstellen von $\alpha_{n-2}(x)$ legen im algebraischen Sinn die $(n-2)$ isotropen Asymptoten von $k^{(n)}$ fest. Es gilt der

SATZ 5. *Der Koeffizient b_n des höchsten Gliedes in der Gleichung (20) eines normierten Tridens $k^{(n)}$ besitzt geometrische Bedeutung bezüglich der isotropen Bewegungsgruppe $\mathfrak{B}_3$. Ist P ein Punkt, der weder dem Tridens $k^{(n)}$, noch der Asymptotenmenge des Tridens angehört, und bezeichnet $\tilde{P}$ den Schnittpunkt der isotropen Geraden durch P mit $k^{(n)}$, während $\mathfrak{k}(\alpha(P))$ die Potenz von P bezüglich der Asymptotenmenge von $k^{(n)}$ bedeutet, so gilt*

$$b_n = \frac{\mathfrak{k}(\alpha(P))\cdot s(\tilde{P}, P)}{\mathfrak{k}(P)}. \tag{21}$$

BEWEIS. Die Asymptotenmenge $\alpha_{n-2}(x)=x^{n-2}+\ldots+a_0=0$ des normierten Tridens $k^{(n)}$ ist ebenfalls eine vollständig zirkuläre Kurve $(n-2)$-ter Ordnung und nach (7) gilt $\mathfrak{k}(\alpha(\mathbf{P}))=(-1)^{n-2}\alpha_{n-2}(x_0)$, wenn $\mathbf{P}$ die Koordinaten $P(x_0, y_0)$ besitzt. Weiter findet man $\tilde{\mathbf{P}}\left(x_0, -\dfrac{\beta_n(x_0)}{\alpha_{n-2}(x_0)}\right)$ und hieraus $s(\tilde{\mathbf{P}}, \mathbf{P})=\dfrac{1}{\alpha_{n-2}(x_0)}[\beta_n(x_0)+$ $+y_0\alpha_{n-2}(x_0)]$, womit wegen $\mathfrak{k}(\mathbf{P})=\dfrac{(-1)^n}{b^n}[\alpha_{n-2}(x_0)y_0+\beta_n(x_0)]$ die Behauptung folgt w.z.z.w.

Der Satz 5 liefert für $n=3$ eine bisher ausstehende geometrische Deutung des Koeffizienten a in (1), während der Satz 4 für $n=3$ die entsprechende Deutung des Koeffizienten $\mathfrak{a}$ in (15) liefert. Daß a der einzige wesentliche Koeffizient in (15) ist, hat K. Strubecker in [5, 143] gezeigt. Wir deuten noch den Koeffizienten a in der Gleichung (14) der divergenten Parabeln im Sinne der Gruppe $\mathfrak{B}_3$:

Satz 6. *Ist P eine Punkt, der nicht auf einer divergenten Parabel $p^{(3)}$ liegt und bezeichnen $\tilde{P}_1$ und $\tilde{P}_2$ die Schnittpunkte der isotropen Geraden durch P mit $p^{(3)}$, dann gilt für den Koeffizienten a des höchsten Gliedes in der Gleichung* (14) *von $p^{(3)}$*

$$a=\frac{s(P,\tilde{P}_1)\cdot s(P,\tilde{P}_2)}{\mathfrak{k}(P)}. \tag{22}$$

Der *Beweis* dieses Satzes ergibt sich unschwer aus (14) unter Anwendung von (7).

Der folgende Satz verallgemeinert ein Resultat aus [2] und liefert weiter einen interessanten *Reduktionssatz* für Tridens-Kurven $k^{(n)}$ mit nur reellen Asymptoten.

Satz 7. *Ist $k^{(n)}$ ein Tridens n-ter Ordnung und l eine isotrope Gerade, so liegen alle Punkte P, für die das Produkt der Potenzen bezüglich $k^{(n)}$ und l konstant ist, auf einem Tridens $k^{(n+1)}$ der Ordnung $(n+1)$, welcher l und die Asymptoten von $k^{(n)}$ als Asymptoten besitzt. Jeder Tridens $k^{(n)}$ mit nur reellen Asymptoten läßt sich durch eine Kette von Operationen, wie sie im ersten Teil des Satzes genannt wurden, aus einem isotropen Kreis erzeugen.*

Beweis. (1) Wird l durch $x=c$ beschrieben und $k^{(n)}$ durch (20) festgelegt, so folgt mit einer Konstanten γ die Bedingung $[\alpha_{n-2}(x)y+\beta_n(x)]\cdot[x-c]=(-1)^{n+1}\gamma b_n$, d. h. man findet

$$[\alpha_{n-2}(x)(x-c)]y+\beta_n(x)(x-c)-(-1)^{n+1}\gamma b_n = 0. \tag{23}$$

Dies ist ein verallgemeinerter, normierter Tridens der Ordnung $(n+1)$ mit den angegebenen Asymptoten.

(2) Ist ein normierter Tridens $k^{(n+1)}$ der Gestalt $\alpha_{n-1}(x)y+\beta_{n+1}(x)=0$ vorgegeben, so bleiben, um eine Reduktion um 1 vorzunehmen, die Gleichungen

$$\begin{aligned}\alpha_{n-1}(x) &= \alpha_{n-2}(x)(x-c)\\ \beta_{n+1}(x) &= \beta_n(x)(x-c)-(-1)^{n+1}\gamma b_n\end{aligned} \tag{24a, b}$$

zu studieren. Nach Voraussetzung besitzt $\alpha_{n-1}(x)$ reelle Nullstellen. Bezeichnet $x=c$ eine solche, so ist wegen der Normiertheit $\alpha_{n-2}(x)$ eindeutig bestimmt. Nun folgt aus (24.b) $\beta_n(x)=\dfrac{\beta_{n+1}(x)+(-1)^{n+1}\gamma b_n}{(x-c)}$, wobei notwendig $b_{n+1}=b_n$ gilt. Wird nun γ so gewählt, daß $x=c$ eine Nullstelle von $\beta_{n+1}(x)+(-1)^{n+1}\gamma b_n$ wird, so liegt auch $\beta_n(x)$ eindeutig fest. Mit der Geraden $x=c$ und der Konstanten γ kann somit der Tridens $k^{(n+1)}$ aus einem Tridens $k^{(n)}$ nach dem angegebenen Verfahren erzeugt werden. Durch sukzessive Anwendung dieses Verfahrens gelangt man schließlich zu einem isotropen Kreis und der Asymptotenmenge

w.z.z.w.

Besitzt $k^{(n+1)}$ auch konjugiert-komplexe Asymptoten, so läßt sich ein entsprechender eleganter Reduktionssatz offenbar nicht angeben.

LITERATURVERZEICHNIS

[1] Burau, W., *Algebraische Kurven und Flächen. Bd. I. Algebraische Kurven der Ebene*, Sammlung Göschen, Bd. **435**, Walter de Gruyter, Berlin, 1962. *MR* **27** # 2899.

[2] Palman, D., Über zirkuläre Kurven 3. Ordnung der isotropen Ebene, Rad JAZU 1987 (im Druck).

[2a] Palman, D., Vollständig zirkuläre Kurven 4. Ordnung der isotropen Ebene, Rad JAZU 1989 (im Druck).

[3] Sachs, H., *Lehrbuch der ebenen isotropen Geometrie*, Vieweg-Verlag, Wiesbaden, 1987.

[4] Strubecker, K., Geometrie in einer isotropen Ebene, *Math.-Naturwiss. Unterricht* **15** (1962), 297—306, 343—351, 385—394. *MR* **26** # 5465a, 5465b; **27** # 631.

[5] Strubecker, K., Über die Parabeln zweiter bis vierter Ordnung, *Praxis d. Math.* **4** (1962), 141—144, 169—174, 197—201.

[6] Strubecker, K., Äquiforme Geometrie der isotropen Ebene, *Arch. Math.* **3** (1952), 145—153. *MR* **14**—405.

[7] Strubecker, K., Zwei Anwendungen der isotropen Dreiecksgeometrie auf ebene Ausgleichsprobleme, *Sitzber. Akad. Wiss. Wien* **192** (1983), 497—559.

[8] Wills, J. M. and Schulte, E., Kepler—Poinsot-type realization of regular maps of Klein, Gordan and Sherk, *Mathem. Forschungsberichte der Universität Siegen*, Nr. **146** (1985).

(Eingegangen am 29. Juni 1987)

MATHEMATISCHES INSTITUT
MONTANUNIVERSITÄT LEOBEN
FRANZ-JOSEF-STRASSE 18
A—8700 LEOBEN
AUSTRIA

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 385—389

ON A PROBLEM OF VAN DOUWEN

I. JUHÁSZ

Abstract

Let $D(\varkappa)$ denote the statement "every first countable T_1 space in which two disjoint closed sets of size $<\varkappa$ can always be separated is normal". We prove here that (i) if $\varkappa$ is regular and $E(\varkappa)$ holds then $D(\varkappa)$ fails; (ii) if $\varkappa$ is strongly compact then $D(\varkappa)$ holds; (iii) if we add $\varkappa$ many Cohen or random reals to V, where $\varkappa$ is strongly compact, then $D(c^+)$ holds in the extension. This shows, in response to a problem of van Douwen [1] that $D(c^+)$ is independent.

In [1] E. van Douwen raised the following problem: Let X be a first countable T_1 space such that any pair of disjoint closed subsets of X of cardinality $\leqq c = 2^{\aleph_0}$ can be separated. Is it true then that X is normal?

The aim of this note is to show that, modulo some large cardinals, the answer to van Douwen's question is independent of set theory. In order to facilitate the formulation of our results we introduce the following piece of notation:

$D(\varkappa)$ is the statement "Every first countable T_1 space X is normal provided that any pair of disjoint closed subsets of X of cardinality $<\varkappa$ can be separated". Thus van Douwen's question may now be phrased as follows: Is $D(c^+)$ true?

We first give a negative answer to this question, which will be a corollary to the next result.

THEOREM 1. *Let $\varkappa$ be an uncountable regular cardinal for which $E(\varkappa)$ holds; then there exists a non-normal but 0-dimensional T_2 and first countable space X which is a $\varkappa$-type increasing union of clopen metrizable subspaces.*

PROOF. Let us start by recalling (cf. [5]) that $E(\varkappa)$ means the existence of a set $E\subset\varkappa$ consisting of ω-limits which is stationary in $\varkappa$, but for each $\alpha\in\varkappa$ the set $E\cap\alpha$ is non-stationary in α.

Now, for every $\delta\in E$ we fix an ω-sequence $\delta_n\in\delta$ that converges up to δ. The underlying set of our space X will be $E\times(\omega+1)$, every point in $E\times\omega$ will be isolated, finally a neighbourhood base for a point $\langle\delta,\omega\rangle\in E\times\{\omega\}$ is formed by the sets

$$V_n(\delta) = \{\langle\delta,\omega\rangle\}\cup\bigcup\{(\delta\setminus\delta_k)\cap E\times\{k\}\colon k\in\omega\setminus n\}$$

for $n\in\omega$.

It is straightforward to check that every $V_n(\delta)$ is clopen, hence X is 0-dimensional, first countable, and trivially T_2.

Research supported by Hungarian National Foundation for Scientific Research Grant no. 1805.

1980 *Mathematics Subject Classification* (1985 Revision). Primary 03E35, 54A35; Secondary 54D15.

Key words and phrases. Normal space, lynx.

To show that X is not normal, consider first two disjoint subsets H and K of E that are both stationary in $\varkappa$, see e.g. [7]. Then $H\times\{\omega\}$ and $K\times\{\omega\}$ are disjoint closed sets in X that we claim cannot be separated. Indeed, let $G\supset H\times\{\omega\}$ be any open set in X, then for every $\delta\in H$ there is some $n\in\omega$ for which $V_n(\delta)\subset G$. Now there is an $m\in\omega$ and a stationary set $S\subset H$ such that $V_m(\delta)\subset G$ for each $\delta\in S$. For every $n\geqq m$ the map φ_n which assigns to every $\delta\in S$ the ordinal δ_n is clearly regressive on S, hence by the pressing down lemma there is a stationary set $S_n\subset S$ and an ordinal $\gamma_n\in K$ such that for every $\delta\in S_n$ we have $\varphi_n(\delta)=\gamma_n$. But this clearly implies that, for $\gamma=\bigcup\{\gamma_n\colon n\in\omega\setminus m\}$,

$$(E\setminus\gamma)\times(\omega\setminus m)\subset G.$$

Of course, a similar argument shows that if G_1 is any open set containing $K\times\{\omega\}$ then there are $\gamma_1\in\varkappa$ and $m_1\in\omega$ such that

$$(E\setminus\gamma_1)\times(\omega\setminus m_1)\subset G_1,$$

hence clearly $G\cap G_1\neq\emptyset$.

Finally, to show the last mentioned property of X it will obviously suffice to prove that for every $\alpha\in\varkappa$ the subspace

$$Y_\alpha=(E\cap\alpha)\times(\omega+1),$$

that is clearly clopen in X if E has a largest element below α, is metrizable. We do this by induction on α:

First, if α is limit (and Y_β has already been shown to be metrizable for all $\beta<\alpha$) then we can choose a closed unbounded set C_α in α such that $C_\alpha\cap E=\emptyset$, i.e. $E\cap\alpha\subset\subset\alpha-C_\alpha$. Let $C_\alpha=\{\beta_\xi\colon \xi\in\varphi\}$ be the increasing enumeration of C_α. We can now write

$$Y_\alpha=\bigcup\{[(\beta_\xi,\beta_{\xi+1})\cap E]\times(\omega+1)\colon \xi\in\varphi\},$$

hence, by the inductive hypothesis, Y_α is metrizable, being the topological sum of metrizable subspaces.

Next, if $\alpha=\delta+1$ with $\delta\in E$ (if $\delta\notin E$ then $Y_\alpha=Y_\delta$), then first note that $Y_\alpha\setminus\{\langle\delta,\omega\rangle\}=Y_\delta\cup(\{\delta\}\times\omega)$ is metrizable if Y_δ is, moreover, since Y_α is regular and first countable, Lemma 2 of [5] implies that Y_α is also metrizable.

COROLLARY. *If $\varkappa>\omega$ is regular and satisfies $E(\varkappa)$ then $D(\varkappa)$ fails. In particular, if there is a regular cardinal $\varkappa>\mathfrak{c}$ with $E(\varkappa)$ then $D(\mathfrak{c}^+)$ fails.*

It is known that the failure of the assumption in the second part of this corollary implies the consistency of certain large cardinals (e.g. many measurables) and in fact it is "very close" to being equiconsistent with the existence of a strongly compact cardinal (cf. e.g. [3]). Therefore the occurence of a strongly compact cardinal in our following results is very natural.

Let us now give here a definition taken from [4]: A family of sets $\mathcal{H}$ in a space X is called $\varkappa$-separated if for every set $Z\in[X]^{<\varkappa}$ the trace family

$$\mathcal{H}\restriction Z=\{H\cap Z\colon H\in\mathcal{H}\}$$

is separated. We denote by $S(\varkappa)$ the following statement: "Every $\varkappa$-separated family in a first countable space is also separated." The relevance of these concepts to our problem will be clear from the next result.

LEMMA 1. *If $\varkappa$ is ω-inaccessible (i.e. for every $\lambda<\varkappa$ we have $\lambda^\omega<\varkappa$) and $S(\varkappa)$ holds then $D(\varkappa)$ is valid.*

PROOF. Let X be a first countable T_1 space in which pairs of disjoint closed sets of size $<\varkappa$ can always be separated (note that in this case X is automatically T_2 as well) and A, B be two disjoint closed sets in X. By $S(\varkappa)$, it will suffice to show that the two-element family $\{A, B\}$ is $\varkappa$-separated in X. But for every $Z\in[X]^{<\varkappa}$ note that $|\overline{A\cap Z}|\leqq|A\cap Z|^\omega<\varkappa$ and similarly we have $|\overline{B\cap Z}|<\varkappa$, since X is T_2 and $\varkappa$ is ω-inaccessible, hence, by assumption, even $\overline{A\cap Z}$ and $\overline{B\cap Z}$ may be separated.

In [4] general conditions under which $S(\varkappa)$ is valid are formulated (in part explicitely, in part just implicitely), however no proofs of these are given there. But since we need them for concrete applications, we decided to present here the details. In fact, for reasons of presentation we start with an easier particular case.

THEOREM 2. *If $\varkappa$ is strongly compact then $D(\varkappa)$ is valid.*

PROOF. Now $\varkappa$ is inaccessible, hence also ω-inaccessible, and thus to apply Lemma 1 we only need to prove $S(\varkappa)$. To this end, let X be first countable (for each $p\in X$ we fix a countable neighbourhood base $\mathscr{V}_p$) and let $\mathscr{U}$ be a fine ultrafilter on $P([X]^{<\varkappa})$ (see e.g. [6]), i.e. such that $\mathscr{U}$ is $\varkappa$-complete and for every $p\in X$

$$\mathscr{Y}_p=\{Z\in[X]^{<\varkappa}\colon p\in Z\}\in\mathscr{U}.$$

Now, suppose that $\mathscr{H}\subset P(X)$ is $\varkappa$-separated, i.e. for every $Z\in[X]^{<\varkappa}$ we have a disjoint open family

$$\mathscr{V}_Z=\{V(H,Z)\colon H\in\mathscr{H}\}$$

with $H\cap Z\subset V(H,Z)$ for $H\in\mathscr{H}$.

For each $Z\in[X]^{<\varkappa}$ and every point $p\in Z\cap\bigcup\mathscr{H}$ we may then choose a neighbourhood $W(p,Z)\in\mathscr{V}_p$ such that

$$W(p,Z)\subset V(H,Z)$$

for $p\in H$.

Now, for any fixed $p\in\bigcup\mathscr{H}$ we have

$$\mathscr{Y}_p=\bigcup\{\mathscr{Y}_{p,V}\colon V\in\mathscr{V}_p\},$$

where

$$\mathscr{Y}_{p,V}=\{Z\in\mathscr{Y}_p\colon W(p,Z)=V\}.$$

Since $\mathscr{Y}_p\in\mathscr{U}$ and $\mathscr{U}$ is $\varkappa$-complete, there is for every $p\in\bigcup\mathscr{H}$ a $V_p\in\mathscr{V}_p$ such that $\mathscr{Y}^p=\mathscr{Y}_{p,V_p}\in\mathscr{U}$ as well. Let us put for each $H\in\mathscr{H}$

$$U_H=\bigcup\{V_p\colon p\in H\}.$$

We claim that the family $\{U_H\colon H\in\mathscr{H}\}$ is disjoint, hence separates $\mathscr{H}$.

To see this, let H_1, H_2 be distinct members of $\mathscr{H}$ and let $p_1\in H_1$ and $p_2\in H_2$; we first choose any $Z\in\mathscr{Y}^{p_1}\cap\mathscr{Y}^{p_2}$. Then, by definition, we have

$$V_{p_1}\subset V(H_1,Z)\quad\text{and}\quad V_{p_2}\subset V(H_2,Z),$$

hence $V_{p_1}\cap V_{p_2}=\emptyset$. This completes the proof.

Now, in order to formulate Fleissner's general result on when $S(\varkappa)$ is valid we need a definition that is also due to him. A pair $\langle B, L\rangle$ is said to be a *lynx* if B is a Boolean algebra, $L\subset B$ is linked, i.e. $b_1, b_2\in L$ implies $b_1\wedge b_2\neq 0$, moreover for every collection $\{b_n: n\in\omega\}\subset B$ with $\bigvee\{b_n: n\in\omega\}=1$ there is a $k\in\omega$ such that $b_0\vee\ldots\vee b_k\in L$. Every measure algebra is trivially a lynx, moreover, by the so-called Dow's lemma [2], so is every regular open algebra $RO(2^\lambda)$ for every λ.

LEMMA 2 (Fleissner). *Assume that for every X with $|X|\geqq\varkappa$ there is a countably complete Boolean homomorphism*

$$h: P([X]^{<\varkappa}) \to B$$

where $\langle B, L\rangle$ is a lynx and $h(\mathcal{Y}_p)=1$ for each $p\in X$ (here, as above, $\mathcal{Y}_p=\{Z\in[X]^{<\varkappa}: p\in Z\}$). Then $S(\varkappa)$ is valid.

PROOF. Let $\mathcal{H}$ be a $\varkappa$-separated family in the first countable space X, and pick h, B, L as above. Just as in the proof of Theorem 2 we get to the equality

$$\mathcal{Y}_p = \bigcup\{\mathcal{Y}_{p,V}: V\in\mathcal{V}_p\}$$

for each $p\in\bigcup\mathcal{H}$. Then using the countable completeness of h we get

$$\bigvee\{h(\mathcal{Y}_{p,V}): V\in\mathcal{V}_p\} = 1.$$

We may then choose $\mathcal{U}_p\in[\mathcal{V}_p]^{<\omega}$ such that

$$\bigvee\{h(\mathcal{Y}_{p,V}: V\in\mathcal{U}_p)\} = h(\bigcup\{\mathcal{Y}_{p,V}: V\in\mathcal{U}_p\})\in L.$$

Clearly, $V_p=\bigcap\mathcal{U}_p$ is a neighbourhood of p, hence

$$U_H = \bigcup\{V_p: p\in H\}$$

is again a neighbourhood of H for each $H\in\mathcal{H}$. To show that $\{U_H: H\in\mathcal{H}\}$ separates $\mathcal{H}$, consider $H_1\neq H_2$ from $\mathcal{H}$ and $p_1\in H_1$, $p_2\in H_2$. Then the linkedness of L implies the existence of $V_1\in\mathcal{U}_{p_1}$ and $V_2\in\mathcal{U}_{p_2}$ such that

$$\mathcal{Y}_{p_1,V_1}\cap\mathcal{Y}_{p_2,V_2}\neq\emptyset,$$

hence let $Z\in\mathcal{Y}_{p_1,V_1}\cap\mathcal{Y}_{p_2,V_2}$. Then

$$V_{p_i}\subset V_i\subset V(H_i, Z)$$

for $i=1, 2$, consequently $V_{p_1}\cap V_{p_2}=\emptyset$, showing that the collection $\{U_H: H\in\mathcal{H}\}$ is disjoint.

Now, it follows from Theorem 6.1 of [4] that the assumption of Lemma 2 is valid in $V^{P_\varkappa}$ if $\varkappa$ is strongly compact and $P_\varkappa$ is the notion of forcing that adds $\varkappa$ many Cohen or random reals to V. Consequently, we obtain the following result.

THEOREM 3. *If $\varkappa$ is strongly compact and $P_\varkappa$ adds $\varkappa$-many Cohen or random reals to V then, in $V^{P_\varkappa}$, we have $D(\varkappa^+)=D(c^+)$.*

PROOF. By the above we have $S(\varkappa)$ and thus also $S(\varkappa^+)=S(c^+)$ in $V^{P_\varkappa}$. Moreover, since c^+ is clearly ω-inaccessible we get $D(c^+)$ from Lemma 1.

Let us remark here that our constructions also yield the independence of the following statement analogous to van Douwen's original: A first countable T_1 space is collectionwise normal if every family of at most (or less than) c closed sets of size $\leqq c$ is separated in X.

In [8] a non-normal, first countable, locally compact T_2 space X of cardinality c is constructed, in ZFC, with the additional property that every closed set in X is either countable or of cardinality c. Clearly, this space X shows that $D(c)$ is simply false, thus, in general, $S(\varkappa)$ does not imply $D(\varkappa)$.

Finally we want to mention a problem that is left open by our above results, namely could $D(c^+)$ be valid when c is small? More precisely we state the following

PROBLEM. Can $D(\aleph_2)$ be consistent? Or, in other words, does CH imply that $D(\aleph_2)$ is false?

REFERENCES

[1] DOUWEN, E. K. VAN, Problem A. 19, *Top. Proc.* **9** (1984), 367.
[2] DOW, A., Remote points in large products, *Topology Appl.* **16** (1983), 11—17. *MR* **85d**: 54027.
[3] FLEISSNER, W. G., The normal Moore space conjecture..., in: *Handbook of set-theoretic topology*, North-Hol, Amsterdam.
[4] FLEISSNER, W. G., Homomorphism axioms and lynxes, in: *Axiomatic set theory*, AMS Contemporary Math. Series, Vol. 31, 1984, 79—97.
[5] HAJNAL, A. and JUHÁSZ I., On spaces in which every small subspace is metrizable, *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.* **24** (1976), 727—731. *MR* **56** #6619.
[6] KANAMORI, A. and MAGIDOR, M., The evolution of large cardinal axioms in set theory, *Higher set theory* (Proc. Conf., Math. Forschungsinst., Oberwolfach, 1977), Lecture notes in math., Vol. 669, Springer, Berlin, 1978, 99—275. *MR* **80b**: 03083.
[7] SOLOVAY, R., Real valued measurable cardinals, *Axiomatic set theory* (Proc. Sympos. Pure Math. Vol. XIII, Part I, Univ. California, Los Angeles, Calif., 1967), Amer. Math. Soc., Providence, R. I., 1971, 397—428. *MR* **45** #55.
[8] WEISS, W., A countably paracompact nonnormal space, *Proc. Amer. Math. Soc.* **79** (1980), 487—490. *MR* **81f**: 54021.

(Received July 31, 1987)

MTA MATEMATIKAI KUTATÓ INTÉZETE
POSTAFIÓK 127
H—1364 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 391—401

ORDER OF MEAN CONVERGENCE OF HERMITE—FEJÉR INTERPOLATION

PÉTER VÉRTESI* and YUAN XU

1. Introduction. Notations. Preliminary results

1.1. The main purpose of this paper is to investigate the order of weighted L^p, $0<p<\infty$, convergence of Hermite—Fejér (HF, in short) interpolating processes based on the roots of generalized Jacobi polynomials (c.f. Theorem 2.1). Moreover, we obtain the corresponding statement concerning Lagrange interpolation, too (c.f. Statement 2.4).

1.2. If x_{kn}, $k=1, 2, ..., n$, are n distinct numbers and f is a bounded function, then the HF interpolatory polynomial is defined to be the unique polynomial of degree at most $2n-1$ satisfying

$$(1.1) \qquad H_n(f, x_{kn}) = f(x_{kn}), \quad H_n'(f, x_{kn}) = 0, \quad k = 1, 2, ..., n.$$

Although there are many-many papers dealing with the convergence and divergence of HF interpolation, they are concerned mostly with uniform norm and, as far as we know, only two papers investigate L^p convergence. In [1], the first of us in a joint work with P. Nevai gives necessary and sufficient conditions for weighted L^p convergence of HF interpolation based on the zeros of generalized Jacobi polynomial (Theorem 1.1). The other one, a very recent paper by J. Prasad and A. K. Varma [2] considers, among others, the order of L^p convergence when the HF interpolation is based on the Tchebycheff roots $\{\cos(2k-1)\pi/(2n)\}$ (Theorem 1.3).

1.3. As it was mentioned, we also deal with the order of weighted L^p convergence of HF interpolation, but we consider the roots of generalized Jacobi polynomials. To state the corresponding results, we introduce the following notations.

N denotes the set of positive integers. The symbol "const" denotes some constant which is positive and independent of the variables and indices. Whenever "const" is used it will always be clear what variables and indices it is independent of. In each formula "const" may take a different value. The symbol "$\sim$" is used as follows. If A and B are two expressions depending on some variables and indices then

$$A \sim B \Leftrightarrow |AB^{-1}| \leqq \text{const} \quad \text{and} \quad |A^{-1}B| \leqq \text{const}.$$

* The work was completed during this author's visit at Temple University, Philadelphia, PA, USA in January—May, 1987, it was partially supported by Hungarian National Foundation for Scientific Research Grant #1801.

1980 *Mathematics Subject Classification* (1985 Revision). Primary 41A05.
Key words and phrases. Hermite—Fejér interpolation, mean convergence, Jacobi polynomials.

Orthogonal polynomials. Let w be a nonnegative Lebesgue-integrable function in $[-1, 1]$ such that

$$\int_{-1}^{1} w > 0.$$

The corresponding set of orthogonal polynomials is denoted by $\{p_n(w)\}$:

$$p_n(w, x) = \gamma_n(w)\, x^n + \text{lower degree terms}, \quad \gamma_n(w) > 0 \quad \text{and}$$

$$\int_{-1}^{1} p_n(w)\, p_m(w)\, w = \delta_{nm}.$$

The zeros of $p_n(w)$ are denoted by $x_{kn}(w)$ and they are indexed so that

$$x_{1n}(w) > x_{2n}(w) > \ldots > x_{nn}(w).$$

The reproducing kernel $K_n(w)$ is defined by

$$K_n(w, x, t) = \sum_{k=0}^{n-1} p_k(w, x)\, p_k(w, t).$$

According to the Christoffel—Darboux formula [3, p. 43] $K_n(w)$ can be written as

$$K_n(x, t) = \frac{\gamma_{n-1}(w)}{\gamma_n(w)} [p_n(w, x)\, p_{n-1}(w, t) - p_{n-1}(w, x)\, p_n(w, t)](x-t)^{-1}.$$

The Christoffel function $\lambda_n(w, x)$ is defined by

$$\lambda_n(w, x)^{-1} = K_n(w, x, x).$$

It is well-known [3, Theorem 3.1.3] that

$$\lambda_n(w, x) = \min \int_{-1}^{1} |P(t)|^2\, w(t)\, dt \tag{1.2}$$

where the minimum is taken over all polynomials P of degree less than n such that $P(x) = 1$. The numbers $\lambda_{kn}(w)$ defined by

$$\lambda_{kn}(w) = \lambda_n(w, x_{kn}(w))$$

are called the Cotes numbers. By the Gauss—Jacobi quadrature formula [3, p. 47]

$$\sum_{k=1}^{n} P(x_{kn}(w))\, \lambda_{kn}(w) = \int_{-1}^{1} Pw$$

holds for every polynomial P of degree less than $2n$. By Szegő's theorem [3, p. 309]

$$0 < \lim_{n\to\infty} \gamma_n(w)\, 2^{-n} < \infty. \tag{1.3}$$

Lagrange interpolation. The Lagrange interpolating polynomials corresponding to the distribution w and bounded function f are denoted by $L_n(w, f)$. They satisfy

$$L_n(w, f, x_{kn}(w)) = f(x_{kn}(w)), \quad n \in N, \quad 1 \leq k \leq n.$$

The polynomial $L_n(w,f)$ can conveniently be written in the form

$$L_n(w,f) = \sum_{k=1}^{n} f(x_{kn}(w))\, l_{kn}(w) \tag{1.4}$$

where the fundamental polynomials $l_{kn}(w)$ are defined by

$$l_{kn}(w,x) = \frac{p_n(w,x)}{p_p'(w,x_{kn}(w))(x-x_{kn}(w))}, \quad 1 \leqq k \leqq n.$$

It is well-known [3, p. 48] that

$$l_{kn}(w,x) = \frac{\gamma_{n-1}(w)}{\gamma_n(w)}\, \lambda_{kn}(w)\, p_{n-1}(w,x_{kn}(w))\, \frac{p_n(w,x)}{x-x_{kn}(w)}. \tag{1.5}$$

Hermite—Fejér interpolation. If the interpolation nodes $\{x_{kn}\}$ in (1.1) are taken to be the zeros $\{x_{kn}(w)\}$ of the orthogonal polynomials $p_n(w)$, then we denote the corresponding Hermite—Fejér interpolating polynomial by $H_n(w,x)$. Hence

$$H_n(w,f,x) = \sum_{k=1}^{n} f(x_{kn}(w))(1-\frac{p_n'(w,x_{kn}(w))}{p_n''(w,x_{kn}(w))}(x-x_{kn}(x))\, l_{kn}(w,x)^2. \tag{1.6}$$

If P is a polynomial of degree less than $2n$ then

$$P(x) = H_n(w,P,x) + \sum_{k=1}^{n} P'(x_{kn}(w))(x-x_{kn}(w))\, l_{kn}(w,x)^2 \tag{1.7}$$

which is the Hermite interpolation formula [3, p. 331].

L^p and L^p_μ spaces. If $0<p<\infty$ then $f\in L^p$ if $\|f\|_p<\infty$ where

$$\|f\|_p = \Big[\int_{-1}^{1} |f(t)|\, dt\Big]^{1/p}, \quad 0<p<\infty,$$

and

$$\|f\|_\infty = \operatorname*{ess\,sup}_{t\in[-1,1]} |f(t)|.$$

If $\mu\geqq 0$ and $0<p<\infty$ then $f\in L^p_\mu$ if $\|f\|_{\mu,p}<\infty$ where

$$\|f\|_{\mu,p} = \Big[\int_{-1}^{1} |f(t)|^p \mu(t)dt\Big]^{1/p}.$$

Naturally, when $0<p<1$, $\|\cdot\|_{\mu,p}$ and $\|\cdot\|_p$ are not norms, nevertheless we retain this notation for convenience.

Jacobi weights. The function u is called a *Jacobi weight* function if u can be written the form

$$u(x) = (1-x)^d(1+x)^b \tag{1.8}$$

for $-1\leqq x\leqq 1$ and $u(x)=0$ for $|x|>1$. In this paper we do not necessarily assume that u is integrable.

Generalized Jacobi polynomials. Let w be a nonnegative integrable function defined in $[-1, 1]$. We say that w is a *generalized Jacobi weight* function ($w\in GJ$) if w can be written in the form

$$w(x) = g(x)(1-x)^A(1+x)^B, \quad -1 \leqq x \leqq 1, \tag{1.9}$$

where $A, B > -1$ and $g^{\pm 1}\in L^\infty$. If, in addition, $g(x)$ is continuous and its modulus of continuity $\omega(g, t)$ satisfies

$$\int_0^1 \frac{\omega(g,t)}{t}\,dt < \infty,$$

then w is a *generalized smooth Jacobi weight* ($w\in GSJ$). If, moreover, $g'\in \mathrm{Lip}\,1$ in $[-1, 1]$, then w is a *very smooth generalized Jacobi weight* ($w\in GCJ$).

Orthogonal polynomials corresponding to generalized Jacobi weight functions are *generalized Jacobi polynomials* (see the works V. Badkov [5] and P. Nevai [6]).

1.4. In [1, Theorem 5, p. 55] P. Nevai and P. Vértesi proved as follows (the original statement is slightly stronger).

THEOREM 1.1. *Let* $w\in GCJ$, $p>0$ *and let u be a Jacobi weight function. Then*

$$\text{(i)} \qquad \lim_{n\to\infty} \|H_n(w,f)-f\|_{u,p} = 0, \quad f\in C[-1,1],$$

holds iff

$$\text{(ii)} \qquad w^{-p}u\in L^1.$$

As a simple consequence of Theorem 1.1, we may mention the following special cases. Obviously, if $u(x)=(1-x^2)^\gamma$, $w(x)=g(x)(1-x^2)^\alpha$, $\alpha,\gamma>-1$, then (i) holds iff $\gamma-\alpha p>-1$ or $\gamma+1>\alpha p$, from where we get

COROLLARY 1.2. (1) *If* $-1<\alpha\leqq 0$, *then* (i) *holds for any* $p>0$, (2) *if* $\alpha>0$, *then* (i) *holds iff* $p<\dfrac{\gamma+1}{\alpha}$, *the last inequality being* $p<1+\dfrac{1}{\alpha}$ *when* $\alpha=\gamma$.

Considering the rate of convergence, J. Prasad and A. K. Varma [2, Theorem 1] obtained

THEOREM 1.3. *If* $w(x)=(1-x^2)^{-1/2}$ *and* $p>0$, *then*

$$\|H_n(w,f)-f\|_{w,p} \leqq \text{const.}\,\omega\left(f,\frac{1}{n}\right), \quad f\in C[-1,1].$$

2. Results

2.1. We prove the following

THEOREM 2.1. *Let* $w \in GCJ$. $W(x) := w(x)\sqrt{1-x^2}$, $p>0$ *and let* u *be an integrable Jacobi weight function. If*

(a) $\|H_n(w,f)-f\|_{u,p} \leqq \text{const.}\, \omega\left(f, \frac{1}{n}\right), \quad f \in C[-1,1]$,

(b) $W^{-p} u \in L^1$,

(b*) $W^{-p^*} u \in L^1, \quad p^* < p$,

then (b)⇒(a) *and* (a)⇒(b*).

Again, let us consider some special cases. Let $u(x)=(1-x^2)^\gamma$, $w(x)=g(x)\cdot(1-x^2)^\alpha$, $\alpha, \gamma > -1$. Since $W^{-p}u \in L^1$ iff $\gamma - p\left(\alpha+\frac{1}{2}\right) > -1$, then if $-1<\alpha<-\frac{1}{2}$, $W^{-p}u\in L^1$ for any $\gamma>-1$, $p>0$; and if $\alpha>-\frac{1}{2}$, then $W^{-p}u \in L^1$ iff $p<\dfrac{\gamma+1}{\alpha+\frac{1}{2}}$.

So we conclude

COROLLARY 2.2. *If* $u(x)=(1-x^2)^\gamma$ *and* $w(x)=g(x)(1-x^2)^\alpha$, $g>0$, $g' \in \text{Lip}\, 1$, $\alpha, \gamma > -1$ *then*

(1) *if* $-1<\alpha\leqq -\frac{1}{2}$, (a) *holds true for any* $p>0$;

(2) *if* $\alpha > -\frac{1}{2}$, (a) *holds true if* $p<\dfrac{\gamma+1}{\alpha+\frac{1}{2}}$; *and whenever* (a) *holds true,* $p \leqq \dfrac{\gamma+1}{\alpha+\frac{1}{2}}$, *the lats inequality being* $p \leqq 1+\dfrac{1}{1+2\alpha}$ *if* $\alpha=\gamma$.

2.2. REMARK. *It follows from our proof that in* (a) $\omega\left(f, \frac{1}{n}\right)$ *cannot be replaced by* $\omega_2\left(f, \frac{1}{n}\right)$, *say* (see (3.7)). *So Theorem 2.1, in certain sense, is the best possible one.* Further, it is easy to see that (b)⇒(ii), which means, if we demand more, generally we have to suppose more (see (i) and (a)). Although this sounds very natural, as an interesting "counterexample" we shall see the somewhat surprising fact that the corresponding conditions for Lagrange interpolation are equivalent (c.f. statement 2.4). So in the mean convergence case, Lagrange interpolation is better than the Hermite—Fejér one.

2.3. Namely, in his comprehensive work P. Nevai proved as follows (see [4, Theorem 6, p. 695]; we use a special case only).

THEOREM 2.3. *Let* $w \in GSJ$, $p>0$ *and let* u *be an integrable Jacobi weigth. Then*

$$\lim_{n\to\infty} \|L_n(w,f)-f\|_{u,p} = 0, \quad f \in C[-1,1] \tag{1}$$

holds iff

$$W^{-(1/2)}u \in L^1. \tag{2}$$

Now let $E_n(f)=\min \|f-P\|_\infty$ where the minimum is taken over all polynomials of degree at most n. Then we have (c.f. 3.8)

STATEMENT 2.4. *Let* $w \in GCJ, p>0$ *and let* u *be an integrable Jacobi weight. Then the conditions* (1), (2) *in Theorem 2.3 and*

$$\|L_n(w,f)-f\|_{u,p} \leqq \text{const.}\, E_n(f), \qquad f \in C[-1,1] \tag{3}$$

are equivalent.

As above, we can get the next

COROLLARY 2.5. *If* $u(x)=(1-x^2)^\gamma$ *and* $w(x)=g(x)(1-x^2)^\alpha$, $g>0$, $g' \in \mathrm{Lip}\, 1$, $\alpha, \gamma > -1$, *then*

(a) *if* $-1<\alpha \leqq -\frac{1}{2}$, (3) *holds true for any* $p>0$,

(b) *if* $\alpha > -\frac{1}{2}$, (3) [*or* (2)] *holds true iff* $p<2\dfrac{\gamma+1}{\alpha+\frac{1}{2}}$, *the last inequality being* $p<2+\dfrac{1}{\alpha+\frac{1}{2}}$ *if* $\alpha=\gamma$.

Using [4, Theorem 6, p. 295] we can get the corresponding statements for quasi Lagrange interpolation, too. We omit the details.

3. Proofs

3.1. The proof of Theorem 2.1 is based on the closed connection between $H_n(w,f)$ and $L_n(w,f)$ revealed and used by G. Freud; P. Nevai, P. Vértesi [1]; J. Prasad, A. K. Varma [2]. In their "ad hoc" proof, the latter ones used nice identities valid if $u=w=(1-x^2)^{-1/2}$; for the general setting we applied certain modifications of theorems proved in paper [4] and the statements of [1]. Further, relations (3.6) and (3.7) applied in [2] also, was very useful.

Statement 2.4 is a simple consequence of the modified versions of Theorems 1 and 2 in [4].

3.2. First we summarize some relations on $p_n(w)$ and $\lambda_n(x)$.

Let $w \in GSJ$, and let $x_{kn}(w) = \cos\theta_{kn}$ ($x_{0n}=1$, $x_{n+1,n}=-1$, $0 \leqq \theta_{kn} \leqq \pi$). Then

$$\theta_{k+1,n} - \theta_{kn} \sim \frac{1}{n} \tag{3.1}$$

uniformly for $0 \leqq k \leqq n$, $n \in N$,

$$\lambda_{kn}(w) \sim \frac{1}{n} w(x_{kn}(w)) \sqrt{1 - x_{kn}(w)^2} \tag{3.2}$$

uniformly for $0 \leqq k \leqq n$, $n \in N$,

$$|p_{n-1}(w, x_{kn}(w))| \sim w(x_{kn}(w))^{-1/2}(1 - x_{kn}(w)^2)^{1/4} \tag{3.3}$$

uniformly for $1 \leqq k \leqq n$, $n \in N$,

$$|p_n(w, x)| \leqq \text{const.}\,[w(x)\sqrt{1-x^2}]^{-1/2}, \quad |x| \leqq 1 - \sigma n^{-2} \tag{3.4}$$

uniformly for $n \in N$, where $\sigma > 0$ is a constant, and

$$|p_n(w, x)| \sim \begin{cases} \sqrt{n}\,[w(1-n^{-1})]^{-1/2}, & 1 + x_{1n}(w) \leqq 2x \leqq 2 \\ \sqrt{n}\,[w(-1+n^{-2})]^{-1/2}, & -2 \leqq 2x \leqq -1 + x_{nn}(w), \end{cases} \tag{3.5}$$

uniformly for $n \in N$ (see [1; Lemma 2, p. 35]).

3.3. As mentioned before, we need the following two theorems from [4] and [1].

LEMMA 3.1 [4, Theorem 1, p. 680[1]]. *Let $w \in GSJ$ and $p > 0$. Let v, V be two not necessarily integrable Jacobi weight functions such that $V \in L^p$, $vV \in L^p$, $V/W^{1/2} \in L^p$ $V/W^{1/2} \in L^1$ and $vW^{1/2} \in L^1$, where, as above, $W(x) = w(x)\sqrt{1-x^2}$. Then for any given bounded function f_n, $n \in N$,*

$$\|L_n(w, vf_n)V\|_p \leqq \text{const.}\,\|f_n\|_\infty, \quad n \in N,$$

with some constant independent of $\{f_n\}$.

LEMMA 3.2 [1; Theorem 5, p. 55][1]. *Let $w \in GCJ$, $p > 0$, and let u be Jacobi weight function such that $w^{-1} \in L^p_u$. Then for any given continuous function f_n, $n \in N$,*

$$\|H_n(w, f_n)\|_{u,p} \leqq \text{const.}\,\|f_n\|, \quad n \in N,$$

with some constant independent of $\{f_n\}$.

3.4. By a well-known theorem of S. B. Stečkin [7], there exists a polynomial $R_n(x)$ of degree at most n such that for all x, $-1 \leqq x \leqq 1$,

$$|f(x) - R_n(x)| \leqq \text{const.}\,\omega\left(f, \frac{\sqrt{1-x^2}}{n}\right) \tag{3.6}$$

[1] The original Theorems are stated for a fixed function f. On the other hand, one can see from their proofs that they can be stated in the above mentioned forms, too.

and

$$(1-x^2)^{1/2}\,|R'(x)| \leqq \text{const.}\, n\,\omega\left(f, \frac{1}{n}\right). \tag{3.7}$$

For this $R_n(x)$ we prove as follows.

LEMMA 3.3. *Let* $w \in GCJ$, $W(x)=w(x)\sqrt{1-x^2}$, $p>0$, *and let* u *be an integrable Jacobi weight function. Then*

$$\|H_n(w, R_n)-R_n\|_{u,p} \leqq \text{const.}\, \omega\left(f; \frac{1}{n}\right), \qquad f \in C[-1, 1]$$

holds true if $W^{-p}u \in L^1$.

PROOF. By (1.5) and (1.7), we have

(3.8) $R_n(x)-H_n(w, R_n, x)=$

$$= \frac{\gamma_{n-1}(w)}{\gamma_n(w)} \sum_{k=1}^{n} R_n'(x_{kn})\,\lambda_{kn}(w)\,p_{n-1}(w, x_{kn}(w))\,p_n(w, x)\,l_{kn}(w, x).$$

According to (3.2) and (3.3), we have

$$\begin{aligned}\lambda_{kn}(w)\,p_{n-1}(w, x_{kn}) &= c_{kn}\, w(x_{kn}(w))^{1/2}(1-x_{kn}(w)^2)^{3/4}/n\\ &= c_{kn}\, W(x_{kn}(w))^{1/2}\sqrt{1-x_{kn}(w)^2}/n,\end{aligned}$$

where c_{kn} are constants bounded uniformly for k and n. Define the continuous function $c_n(x)$ such that

$$c_n(x_{kn}(w)) = c_{kn}$$

and $c_n(x)$ uniformly bounded for $-1 \leqq x \leqq 1$ and n. Then if the continuous function $q_n(x)$ is defined by

$$q_n(x) = c_n(x)\, R_n'(x)\sqrt{1-x^2}/n$$

we can rewrite (3.8) as

$$R_n(x)-H_n(w, R_n, x) = \frac{\gamma_{n-1}(w)}{\gamma_n(w)}\, p_n(w, x)\, L_n(w, q_n W^{1/2}, x). \tag{3.9}$$

Furthermore, from (3.7)

$$\|q_n\|_\infty \leqq \text{const.}\, \omega\left(f; \frac{1}{n}\right). \tag{3.10}$$

By Theorem 6.3.14 in [6, p. 113] for every $0<p<\infty$ and Jacobi weight u there exists a constant $\sigma=\sigma(p, u)>0$ such that for every polynomial P of degree at most $2n$

$$\int_{-1}^{1} |p(t)|^p\, u(t)\,dt \leqq 2 \int_{-1+\sigma n^{-2}}^{1-\sigma n^{-2}} |p(t)|^p\, u(t)\,dt.$$

Using this to the polynomial $R_n(x)-H_n(w, R_n, x)$, by (1.4), (3.8) and (3.9), it follows

$$(3.11)\ \|R_n-H_n(w,R_n)\|_{u,p}^p \leqq \text{const.} \int_{-1+\sigma n^{-2}}^{1-\sigma n^{-2}} |p_n(w,x)|^p\, |L_n(w, q_n W^{1/2}, x)|^p\, u(x)\, dx$$

$$\leqq \text{const.} \int_{-1+\sigma n^{-2}}^{1-\sigma n^{-2}} |L_n(w, q_n, W^{1/2}, x)|^p\, W(x)^{-p/2}\, u(x)\, dx$$

$$\leqq \text{const.}\ \|L_n(w, q_n W^{1/2}, x)\, W^{1/2}\|_{u,p}^p.$$

Now we apply Lemma 3.1 with $v=W^{1/2}$, $V=W^{-1/2}u^{1/p}$. First let us check the conditions of Lemma 3.1 which in our special case are $W^{-p/2}u\in L^1$, $u\in L^1$, $W^{-p}u\in L^1$ and $W\in L^1$. Obviously, $W\in L^1$ and $u\in L^1$ (by the conditions of Lemma 3.3). Now, if $w(x)=g(x)(1-x)^a(1+x)^b$, first let $-1<a\leqq -\frac{1}{2}$. Then $W^{-p/2}$ and W^{-p} are integrable on [0, 1]. If $a>-\frac{1}{2}$, then by $W^{-p}u\in L^1$ (condition) $W^{-p/2}u$ is integrable on [0, 1]. The interval $[-1, 0]$ can be treated similarly. So we get from (3.10), (3.11) and Lemma 3.1 that

$$\|R_n-H_n(w,R_n)\|_{u,p} \leqq \text{const.}\ \|q_n\|_\infty \leqq \text{const.}\ \omega\left(f, \frac{1}{n}\right),$$

which was to be proved.

3.6. Proof of Theorem 2.1 (b)$\Rightarrow$(a). Let R_n be a polynomial satisfying (3.6) and (3.7). Then

$$\|H_n(w,f)-f\|_{u,p} \leqq \|f-R_n\|_{u,p}+\|H_n(w,R_n)-R_n\|_{u,p}+\|H_n(w,f-R_n)\|_{u,p}.$$

By (3.6) and Lemma 3.3, we only need consider the third term on the right. Clearly, $W^{-p}u\in L^1$ implies $w^{-1}\in L_u^p$, so (a) follows from Lemma 3.2 and (3.6).

(a)$\Rightarrow$(b*).

Let $f_1(x)=x$. Then $\omega\left(f_1, \frac{1}{n}\right)=\frac{1}{n}$, and

$$f_1(x)-H_n(w,f_1,x)=\frac{\gamma_{n-1}(w)^2}{\gamma_n(w)^2}\, p_n(w,x)^2 \sum_{k=1}^n \frac{\lambda_{kn}(w)^2\, p_{n-1}(x_{kn}(w))^2}{x-x_{kn}(x)}.$$

For $1+x_{1n}\leqq 2x\leqq 2$, by (1.3), (3.2) and (3.3) we have

$$f_1(x)-H_n(w,f_1,x) \geqq \text{const.} \sum_{k=1}^n \lambda_{kn}(w)\, \frac{1-x_{kn}(w)^2}{1-x_{kn}(w)}\, p_n(w,x)^2/n$$

$$\geqq \text{const.} \sum_{k=1}^n \lambda_{kn}(w)\, p_n(w,x)^2/n.$$

Now, by Gauss—Jacobi quadrature formula,

$$\sum_{k=1}^{n} \lambda_{kn}(w) = \int_{-1}^{1} w(x)\,dx = \text{const},$$

$$f_1(x)-H_n(w,f_1,x) \geqq \text{const.}\, p_n(w,x)^2/n \tag{3.12}$$

for $1+x_{1n} \leqq 2x \leqq 2$ (c.f. [1; Lemma 5, p. 43]). By (a) we can write

$$\int_{(1+x_{1n}(w))/2}^{1} |H_n(w,f_1,x)-f_1(x)|^p\, u(x)\,dx \leqq \text{const.}\, n^{-p}$$

which, using (3.12), gives

$$n^{-p} \int_{(1+x_{1n}(w))/2}^{1} |p_n(w,x)|^{2p}\, u(x)\,dx \leqq \text{const.}\, n^{-p}.$$

By (3.5) this is equivalent to

$$w(1-n^{-2})^{-p} \int_{(1+x_{1n}(w))/2}^{1} u(x)\,dx \leqq \text{const.}\, n^{-p}. \tag{3.13}$$

Writing $w(x)=g(x)(1-x)^a(1+x)^b$, $u(x)=(1-x)^A(1+x)^B$ and applying (3.1), we can see that (3.13) is equivalent to

$$ap-A-1+\frac{p}{2} \leqq 0, \quad \text{or} \quad A-ap-\frac{p}{2} \geqq -1, \tag{3.14}$$

from where we get that $W^{-p^*}u$ is integrable on $[0, 1]$ whenever $p^*<p$. The interval $[-1, 0]$ can be treated similarly, which completes the proof of Theorem 2.1.

3.7. By (3.12) and (3.5) we have

$$\|H_n(w,f_1)-f_1\|_{u,p}^p \geqq \int_{(1+x_{1n}(w))/2}^{1} |H_n(w,f_1,x)-f_1(x)|^p u\,(x)\,dx \tag{3.15}$$

$$\geqq \text{const.}\, n^{2(ap-A-1)} > \text{const.}\, \omega_2\left(f_1, \frac{1}{n}\right) = 0,$$

which implies 2.2.

3.8. Proof of Statement 2.4. If $P_n(f)$ is the polynomial for which $\|f-P_n\|_\infty = E_n(f)$, then

$$\|L_n(w,f)-f\|_{u,p} \leqq \|L_n(w,f-P_n)\|_{u,p} + \|f-P_n\|_{u,p} := I_1+I_2.$$

If $W^{-p/2}u \in L^1$, then it is easy to check that Lemma 3.1 can be applied with the cast $f_n=f-P_n$, $V=u$ and $v=1$, which gives that $I_1 \leqq \text{const}\, E_n(f)$, $I_2 \leqq \text{const}\, E_n(f)$ is obvious. So we got (2)⇒(3). By (3)⇒(1) and Theorem 2.3, we get (3)⇒(2). The proof is complete.

REFERENCES

[1] NEVAI, P., VÉRTESI, P., Mean convergence of Hermite—Fejér interpolation. *J. Math. Anal. Appl.* **105** (1986), 26—58.
[2] PRASAD, J. and VARMA, A. K., An analogue of a problem of P. Erdös and E. Feldheim on L_p convergence of interpolating processes (to appear).
[3] SZEGŐ, G., *Orthogonal polynomials*, Amer. Math. Soc. Coll. Publ. Vol. 23, American Mathematical Society, Providence, R. I., 1959. *MR* **21** # 5029.
[4] NEVAI, P., Mean convergence of Lagrange interpolation, III, *Trans. Amer. Math. Soc.* **282** (1984), 669—698. *MR* **85c**: 41009.
[5] BADKOV, V., Convergence in the mean and almost everywhere of Fourier series in polynomials orthogonal on an interval, *Math. USSR Sb.* **24** (1974), 223—256.
[6] NEVAI, P., Orthogonal polynomials, *Mem. Amer. Math. Soc.* **18** (1979), no. 213, Amer. Math. Soc., Providence, R. I., 1979. *MR* **80k**: 42025.
[7] STEČKIN, S. B., Generalization of some Bernstein type inequalities, *Dokl. Akad. Nauk SSSR* (*N. S.*) **60** (1948), 1511—1514 (in Russian). *MR* **9**—579.

(Received September 17, 1987)

MTA MATEMATIKAI KUTATÓ INTÉZETE
POSTAFIÓK 127
H—1364 BUDAPEST
HUNGARY

DEPARTMENT OF MATHEMATICS
TEMPLE UNIVERSITY
PHILADELPHIA, PA 19122
U.S.A.

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 403—406

THIRD NOTE ON HAJNAL—MÁTÉ GRAPHS

PÉTER KOMJÁTH

0. Introduction

This paper is a continuation of [4] and [5]. If G is a graph on ω_1 (only graphs of this kind will be considered), for $\alpha<\omega_1$ $G(\alpha)$ denotes $\{\beta<\alpha\colon \{\beta,\alpha\}\in G\}$. G is a Hajnal—Máté (in short, H—M) graph, if for every $\alpha<\omega_1$ $G(\alpha)$ is either finite or is an ω-sequence, cofinal in α. The *chromatic number*, $\mathrm{Chr}\,(G)$ is the minimal cardinality $\varkappa$ such that there exists an $f\colon \omega_1\to\varkappa$ with $f^{-1}(\{\xi\})$ independent for $\xi<\varkappa$. In [3], A. Hajnal and A. Máté showed that, assuming $\diamondsuit^*$ there exists an H—M graph G with $\mathrm{Chr}\,(G)=\omega_1$. (For the combinatorial principles $\diamondsuit$, $\diamondsuit^*$ and set theoretic background see [7].) In [4] a similar example without triangle is given, and in [5] even those circuits built from two monotonic paths are excluded. A graph Γ is *obligatory* if it is a (not necessarily spanned) subgraph of every G with $\mathrm{Chr}\,(G)\geqq\omega_1$. In [2] the results of [4], [5] were applied to limit (from $\diamondsuit^*$) the obligatory graphs.

Our first theorem gives an H—M graph G with $\mathrm{Chr}\,(G)=\omega_1$, the circuits mentioned above are excluded, and G has the following strange property: there is a function $g\colon [\omega_1]^2\to\omega$ such that if $\alpha<\beta<\omega_1$ are given, every point $\gamma>\beta$, joined to both α and β, is joined to at most $g(\alpha,\beta)$ points between α and β. A corollary of this is that there exists an ω_1-chromatic graph on ω_1 without subgraphs of the following type: the points are $\{x_i, y_s\colon i\leqq\omega,\ s\in[\omega+1]^{<\omega}\}$, $x_0<x_1<\ldots<x_\omega$, x_i is joined to y_s if and only if $i\in s$.

Our last theorem asserts that, under $\diamondsuit^*$ every obligatory graph Γ splits: the vertex set is the union of countable sets $\{A_\alpha\colon \alpha<\omega_1\}$ such that if $x\in A_\alpha$, then x is joined to finitely many points in $\bigcup\{A_\beta\colon \beta<\alpha\}$.

1. Construction of an exotic H—M graph

THEOREM 1 ($\diamondsuit^*$). *There exist a graph G on ω_1 and a function $g\colon [\omega_1]^2\to\omega$ such that*

(a) $\mathrm{Chr}\,(G)=\omega_1$, *$G$ is a Hajnal—Máté graph;*

(b) *no circuit of G is the union of two monotonic paths;*

(c) *if $\alpha<\beta<\xi$ and ξ is joined to both α and β, then it is joined to at most $g(\alpha,\beta)$ points between α and β.*

1980 *Mathematics Subject Classification.* Primary 04A20; Secondary 05C10.
Key words and phrases. Infinite graphs, chromatic number, diamond principle.

PROOF. Let $\mathscr{H}_\alpha \subseteq P(\alpha)$ $(\alpha<\omega_1)$ be the $\diamondsuit^*$-sequence, $\{S_\gamma\colon \gamma<\omega_1\}$ disjoint stationary sets.

We are going to build G and g by transfinite recursion. Assume α is limit and $G|\alpha$, $g|[\alpha]^2$ have already been defined, and $\alpha\in S_\gamma$ with $\gamma<\alpha$.

Put $\mathscr{H}_\alpha=\{A_0, A_1, \dots\}$. Let $\langle \alpha_n\colon n<\omega\rangle$ be a sequence increasingly converging to α. First, define g on $\alpha\times\{\alpha\}$ in such a way that if A_i is infinite, then

$$\sup\{g(\tau,\alpha)\colon \tau\in A_i\}=\omega.$$

Pick those A_i's for which the following is true:

$(*)$ if $s\in[\alpha-\gamma]^{<\omega}$ and $j<\omega$ then

$A_i\cap\{\xi\colon \tau\in s\to g(\tau,\xi)\geqq j$ and ξ is not $\sup(s)$-covered$\}$ is unbounded in α,

where δ is β-covered if and only if there exists a monotonic path $v_0<v_1<\dots<v_n=\delta$ with $v_0\leqq\beta$.

If $A_{i_0}, A_{i_1}, \dots$ is the sequence of the A_i's described above, choose y_t for $t=0, 1, \dots$ with

(a) $\alpha_t<y_t$;

(b) $y_t\in A_{i_t}$;

(c) y_t is not y_{t-1}-covered, y_0 is not γ-covered;

(d) if $t'<t$, then $g(y_{t'}, y_t)\geqq t$.

This is possible by the choice of the A_{i_t}'s. If a forbidden circuit is born with α as last point then for some $t<\omega$ y_t would be y_{t-1}-covered. Clearly, the property of g prescribed in the theorem will be satisfied, too. So we can choose $G(\alpha)=\{y_t\colon t<\omega\}$.

We prove that G is ω_1-chromatic. Assume that $f\colon \omega_1\to\omega$ is a good coloring. Let us call $i<\omega$ *bad* if

$(**)$ for every $\delta<\omega_1$ there are $\delta'<\omega_1$, $j<\omega$, and $s\in[\omega_1-\delta]^{<\omega}$ such that $\{\xi\colon f(\xi)=i,\ \xi$ is not δ'-covered, $\forall\tau\in s\ g(\tau,\xi)\geqq j\}$ is bounded

(in ω_1), say, it is in ε. Otherwise call i good. Now, if i is bad, there is a series of disjoint finite sets s^i_η $(\eta<\omega_1)$ ordinals δ'^i_η, ε^i_η and finite numbers j^i_η of property $(**)$. We can assume that $j^i_\eta=j^i$ $(\eta<\omega)$. By the property $\diamondsuit^*$ for club many α it is true that $B_i\in\mathscr{H}_\alpha$ for every bad i, where $B_i=\bigcup\{s^i_\eta\colon \eta<\omega\}$, moreover $\alpha>\delta'^i_\eta$, ε^i_η $(\eta<\omega)$.

If i is good, there exists a $\gamma_i<\omega_1$ such that for every $s\in[\omega_1-\gamma_i]^{<\omega}$ $\delta'<\omega_1$, $j<\omega$ the set

$$\{\xi\colon f(\xi)=i,\ \ \xi \text{ is not } \delta'\text{-covered},\ \ \forall\tau\in s\ g(\tau,\xi)\geqq j\}$$

is unbounded in ω_1. There exists a set C_i club in ω_1 such that the above statement still holds below α for any $\alpha\in C_i$. Put $\gamma=\sup\{\gamma_i\colon i \text{ good}\}\cup\sup\{\delta'^i_\eta\colon i \text{ bad}, \eta<\omega\}$. Intersecting these C_i's with each other, with a club set D such that if $\alpha\in D$ then for every good i $f^{-1}(i)\cap\alpha\in\mathscr{H}_\alpha$ (by $\diamondsuit^*$), the set described in the previous paragraph and with S_γ, we get a point α which cannot get any of the colors; if $i=f(\alpha)$ is good $i=i_t$ for some $t<\omega$, so $f(\alpha)=f(y_t)$, a contradiction. If $i=f(\alpha)$ is bad, then for some $\eta<\omega$ there is a $\tau\in s^i_\eta$ such that $g(\tau,\alpha)>j^i$. As $\alpha>\varepsilon^i_\eta$ α is δ'^i_η-covered, which is impossible as $\gamma>\delta'^i_\eta$.

2. Excluded subgraphs

THEOREM 2 ($\diamondsuit^*$). *There exists an ω_1-chromatic graph on ω_1 without a subgraph of the following type: the points are $x_0<x_1<\ldots<x_\omega$, y_s $(s\in[\omega+1]^{<\omega})$ such that they are different and x_i is joined to y_s if $i\in s$.*

PROOF. Take the example from Theorem 1. As $j=g(x_0, x_\omega)$ is finite, if $0, \omega\in s$ and $|s|>j$ then $y_s<x_\omega$. If $0, \omega\in s_0, s_1$ and $j<|s_0|, |s_1|$, x_0, y_{s_0}, y_{s_1}, x is a forbidden circuit unless $y_{s_0}, y_{s_1}<x_0$. This gives that x_ω is joined to infinitely many points below x_0, which is impossible, as our graph is an H—M graph.

THEOREM 3 ($\diamondsuit^*$). *Assume that Γ is a graph on ω_1 such that every graph G with $|G|=\mathrm{Chr}\,(G)=\omega_1$ contains Γ. Then there is a closed unbounded set C such that if $c\in C$, $c\leqq x$, then $\Gamma(x)\cap c$ is finite.*

PROOF. Let Γ be as above. We are going to define an H—M graph G with $\mathrm{Chr}\,(G)=\omega_1$. If $f\colon \omega_1\to\omega_1$ is an imbedding of Γ as a subgraph of G, there exists a closed, unbounded set C, such that $\beta<\alpha\to f(\beta), f^{-1}(\beta)<\alpha$ for $\alpha\in C$. As G is H—M, there is at most one point in Γ such that it is joined to infinitely many points in α. Call this point $\beta(\alpha)$, it is defined, once $f|\alpha$ and α are given. We assume that the set $S=\{\alpha\colon \Gamma(\beta(\alpha))\cap\alpha$ is infinite$\}$ is stationary.

By $\diamondsuit^*$, there are functions $f_{\alpha,n}\colon \alpha\to\alpha$ $(n<\omega)$ and sets $X_{\alpha,n}\subseteq\alpha$ $(n<\omega)$ 'catching' the functions $f\colon \omega_1\to\omega_1$ and sets $X\subseteq\omega_1$, respectively. Assume that $G(\beta)$ is defined for $\beta<\alpha$. $G(\alpha)=\emptyset$, unless α is limit. In this latter case, choose distinct points $\{z_1, z_2, \ldots\}$ into G with the z_i's converging to α, $G(\alpha)\cap X_{\alpha,n}\neq\emptyset$ if $X_{\alpha,n}$ is unbounded in α, $G(\alpha)\not\supseteq f''_{\alpha,n}\Gamma(\beta(\alpha))$ where $\beta(\alpha)$ is calculated as $f|\alpha=f_{\alpha,n}$.

To show that G is ω_1-chromatic, assume that $\omega_1=\bigcup\{X_n\colon n<\omega\}$ is a good coloring. There exists an α such that α is limit, if X_n is bounded then $\sup X_n<\alpha$; if X_n is unbounded, then it is unbounded in α, $X_n\cap\alpha$ is an element of our $\diamondsuit^*$-sequence. Clearly, no color can be given to α.

Assume that $f\colon \omega_1\to\omega_1$ imbeds Γ into G. There exists an α such that $f''\alpha, f^{-1''}\alpha\subseteq\alpha$, $\alpha\in S$ (S is calculated from f, but $\alpha\in S$ depends on $f|\alpha, f^{-1}|\alpha$), and there is an $n<\omega$ such that $f|\alpha=f_{\alpha,n}$. As $\alpha\in S$, $\alpha\cap\Gamma(\beta(\alpha))$ is finite, and we cannot find $f(\beta(\alpha))$; it is not α by our construction, and, as G is H—M, it cannot be larger.

REFERENCES

[1] AVRAHAM, U., DEVLIN, K. and SHELAH, S., The consistency with CH of some consequences of Martin's axiom plus $2^{\aleph_0}>\aleph_1$, *Israel J. Math.* **31** (1978), 19—33. *MR* **58** # 21602.

[2] HAJNAL, A. and KOMJÁTH, P., What must and what need not be contained in a graph of uncountable chromatic number?, *Combinatorica* **4** (1984), 47—52. *MR* **85d**: 05115.

[3] HAJNAL, A. and MÁTÉ, A., Set mappings, partitions, and chromatic numbers, *Logic Colloquium '73* (Bristol, 1973), Studies in Logic and the Foundations of Mathematics, **80**, North-Holland, Amsterdam, 1975, 347—379. *MR* **54** # 12528.

[4] KOMJÁTH, P., A note on Hajnal—Máté graphs, *Studia Sci. Math. Hungar.* **15** (1980), 275—276. *MR* **84e**: 03057.

[5] KOMJÁTH, P., A second note on Hajnal—Máté graphs, *Studia Sci. Math. Hungar.* **19** (1984), 245—246.
[6] KOMJÁTH, P., The colouring number, *Proc. London Math. Soc.* **54** (1987), 1—14.
[7] KUNEN, K., *Set theory. An introduction to independence proofs*, Studies in Logic and the Foundations of Mathematics, **102**, North-Holland, Amsterdam, 1980. *MR* **82f**: 03001.

(Received January 20, 1984)

EÖTVÖS LORÁND TUDOMÁNYEGYETEM
TERMÉSZETTUDOMÁNYI KAR
SZÁMÍTÓGÉPTUDOMÁNYI TANSZÉK
MÚZEUM KRT. 6—8
H—1088 BUDAPEST
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 407—428

BOREL COVER AND BOREL EXTENSION

V. K. ZAHAROV

Contents

Introduction

It is well-known which role Borel functions play in some sections of mathematics. They represent probably the most remarkable class among all known extensions of the class of continuous functions.

In spite of the fact that the class of Borel functions is extensively investigated, the natural question, what properties distinguish the *Borel extension* $B^*(T)$ among all the other extensions of the set $C^*(T)$ of all bounded continuous functions on a space T, had no answer.

The paper consists of three paragraphs on three different topics, which are connected by the fact that the proofs of the consecutive results are based on the previous ones.

The first paragraph is devoted to the *Borelian cover* H of a completely regular space T. Briefly, it can be defined as such "good" preimage of the space T that discontinuous Borel functions, lifted on H, become "almost" continuous on H. This cover was considered by Gordon [1], Sentilles [2], [3], Graves [4] and others, however, no characterization of it was known. The first characterization of the Borelian cover was given by the author in the paper [5] with the help of the notion of *perfect preimages lifting separable covering.*

On this base the notion of *extensions of* $C^*(T)$ *inheriting separable decomposition* was introduced. With the help of this notion in the second and the third paragraphs the characterizations of the Borelian extension $B^*(T)$ are presented.

1980 *Mathematics Subject Classification.* Primary 16A86; Secondary 16A08, 54G05, 54C50, 54H05, 46E05.

Key words and phrases. Continuous functions, Borel functions, lattice-ring extensions, vector-lattice extensions, Borelian cover.

In the paper we shall adhere to the terminology accepted in the books [6]—[10].

The author expresses his profound gratitude to Á. Császár and E. Makai Jr. for their help during fulfilling this work at the time of the author's staying at the Budapest University in 1983—84.

1. Perfect preimages lifting separable covering

Let T and H be completely regular spaces and $\eta\colon H\to T$ be a surjective perfect mapping.

1.1. The family of all cozero-sets in H will be denoted by $\mathscr{C}(H)$. Consider some base in H consisting of a subset $\mathscr{C}_0(H)$ of $\mathscr{C}(H)$ and containing H and the empty set. Let $\mathscr{Z}_0(H)$ denote the set of all complements to elements of $\mathscr{C}_0(H)$ and $\Delta_0(H)$ denote the set of all open-closed elements of $\mathscr{C}_0(H)\cap\mathscr{Z}_0(H)$.

The base $\mathscr{C}_0(H)$ will be called *completely normal* if:

1) $\mathscr{C}_0(H)$ is closed under countable unions and finite intersections;

2) disjoint zero-sets from $\mathscr{Z}_0(H)$ are contained in disjoint cozero-sets from $\mathscr{C}_0(H)$;

3) any cozero-set in $\mathscr{C}_0(H)$ is a countable union of zero-sets from $\mathscr{Z}_0(H)$.

Note that if $\mathscr{C}_0(H)$ is a completely normal base then $(H, \mathscr{Z}_0(H))$ is a completely normal Alexandrov space ([11]). Further we shall assume that $\eta^{-1}C\in\mathscr{C}_0(H)$ for any $C\in\mathscr{C}(T)$.

LEMMA. *Let $\mathscr{C}_0(H)$ be a completely normal base. Then*

(a) *for any $C\in\mathscr{C}_0(H)$ there exists a sequence $C_k\in\mathscr{C}_0(H)$ such that $C=\bigcup C_k$ and* $\operatorname{cl} C_k\subset C$;

(b) *for any $s_1\neq s_2\in H$ there exist $C_1, C_2\in\mathscr{C}_0(H)$ such that $s_1\in C_1$, $s_2\in C_2$,* $s_1\notin\operatorname{cl} C_2$, $s_2\notin\operatorname{cl} C_1$ and $C_1\cup C_2=H$;

(c) *for any $C\in\mathscr{C}_0(H)$ and any $s\in C$ there exists $C_1\in\mathscr{C}_0(H)$ such that $s\notin\operatorname{cl} C_1$ and $C\cup C_1=H$.*

PROOF. Consider the set Φ consisting of all functions $f\in C^*(H)$ such that $f^{-1}(]a, b[)\in\mathscr{C}_0(H)$ for any open interval $]a, b[$. Then Φ is a uniformly complete vector-lattice containing the constant functions. By the theorem of Alexandrov $\mathscr{C}_0(H)=\{\operatorname{coz} f\mid f\in\Phi\}$ ([11]). Therefore Φ separates points and closed subsets from H. Now assertions (a)—(c) follow from these facts.

1.2. Let H be a perfect preimage of T with a completely normal base $\mathscr{C}_0(H)$. The preimage H will be called *lower extremally disconnected* if $\operatorname{cl}\eta^{-1}G\in\Delta_0(H)$ for any open set G in T. The preimage H will be called *σ-extremally disconnected* if $\operatorname{cl} C\in\Delta_0(H)$ for any cozero-set $C\in\mathscr{C}_0(H)$.

1.3. Let $\mathscr{S}(T)$ denote the set of all non-empty countable subsets of the space T. For every countable subset S consider the closed separable subset $T_S\equiv\operatorname{cl} S$. The covering $\{T_S\mid S\in\mathscr{S}(T)\}$ will be called the *separable covering of the space T*.

The preimage H will be called *lifting separable covering* if H has a family of closed subsets $\{H_S\mid S\in\mathscr{S}(T)\}$ such that $\bigcup H_S$ is dense in H, $\eta H_S=T_S$ and $S_1\subset S_2$

implies $H_{S_1} \subset H_{S_2}$. The mapping $T_S \mapsto H_S$ will be called the *lifting of the separable covering*.

Let $\{H, \eta\colon H \to T, T_S \mapsto H_S\}$ and $\{\dot{H}, \dot{\eta}\colon \dot{H} \to T, T_S \mapsto \dot{H}_S\}$ be preimages lifting separable covering. The preimage H will be called *larger* than the preimage $\dot{H}$ if there exists a surjective perfect mapping $\gamma\colon H \to \dot{H}$ such that $\eta = \dot{\eta} \circ \gamma$ and $\gamma H_S = \dot{H}_S$. The preimage H will be called *isomorphic* to the preimage $\dot{H}$ if there exist mutually inverse homeomorphisms $\gamma\colon H \to \dot{H}$ and $\delta\colon \dot{H} \to H$ such that H is larger than $\dot{H}$ relative to γ and $\dot{H}$ is larger than H relative to δ.

1.4. Let H be a perfect preimage of T lifting separable covering.

The preimage H will be called *saturated,* if for any H_S and any open set G intersecting H_S there exists an H_R such that $\emptyset \neq H_R \subset H_S \cap G$ and $R \subset S$.

The preimage H will be called *filled* if $\bigcup H_{S_k}$ is dense in H_S for any sequence S_k such that $\bigcup S_k = S$. Any saturated preimage is filled.

The preimage H will be called *lower disjoined* if $\eta^{-1}G \cap H_S = \emptyset$ implies $(\mathrm{cl}\, \eta^{-1}G) \cap H_S = \emptyset$ for any open set G in T.

1.5. Let H be a perfect preimage of T lifting separable covering and having a completely normal base $\mathscr{C}_0(H)$.

The preimage H will be called *σ-disjoined* if $C \cap H_S = \emptyset$ implies $\mathrm{cl}\, C \cap H_S = \emptyset$ for any cozero-set $C \in \mathscr{C}_0(H)$.

2. Vector-lattice extensions inheriting separable decomposition

We shall suppose that all considered vector-lattices are Archimedian, have fixed strong units and are uniformly complete with respect to their units and that all considered vector-lattice homomorphisms preserve these units. Also we shall suppose that all considered vector-lattice ideals are uniformly closed.

Let T be a completely regular space and $C^*(T)$ be the vector-lattice of all bounded continuous functions on T. Let X be a vector-lattice and $u\colon C^*(T) \to X$ be an injective vector-lattice homomorphism. We shall say that X is an *extension* of $C^*(T)$ and shall identify $C^*(T)$ with its image in X.

2.1. Remind that the vector-lattice X is called *σ-Dedekind complete* if for any countable set bounded from above there exists the supremum of this set. The extension X will be called *lower Dedekind complete* if for any set from $C^*(T)$ bounded from above there exists the supremum of this set in X.

Let Y be an ideal in X. The ideal Y is called a *component* of X if $y_\xi \in Y$, $x \in X$ and $x = \sup y_\xi$ imply $x \in Y$. The ideal Y will be called a *σ-component* of X if $y_n \in Y$, $x \in X$ and $x = \sup y_n$ imply $x \in Y$. The ideal Y will be called a *lower component* of X if $y_\xi \in C^*(T) \cap Y$, $x \in X$ and $x = \sup y_\xi$ imply $x \in Y$.

2.2. For any countable set $S \in \mathscr{S}(T)$ consider the ideal

$$C_S^*(T) \equiv \{f \in C^*(T) \mid f(T_S) = 0\}$$

in $C^*(T)$. The family $\{C_S^*(T) \mid S \in \mathscr{S}(T)\}$ will be called the *separable decomposition of the vector-lattice $C^*(T)$.*

The extension X of $C^*(T)$ will be called *inheriting separable decomposition* if X has a family of proper ideals $\{X_S | S \in \mathcal{S}(T)\}$ such that $\cap X_S = \{0\}$, $uf \in X_S$ iff $f \in C_S^*(T)$ and $S_1 \subset S_2$ implies $X_{S_1} \supset X_{S_2}$. The mapping $C_S^*(T) \mapsto X_S$ will be called the *inheritance of separable decomposition.*

Let $\{X, u\colon C^*(T) \to X, C_S^*(T) \mapsto X_S\}$ and $\{\mathring{X}, \mathring{u}\colon C^*(T) \to \mathring{X}, C_S^*(T) \mapsto \mathring{X}_S\}$ be extensions inheriting separable decomposition. The extension X will be called *larger* than the extension $\mathring{X}$ if there exists an injective vector-lattice homomorphism $v\colon \mathring{X} \to X$ such that $v \circ \mathring{u} = u$ and $v\mathring{X}_S \subset X_S$. The extension X will be called *isomorphic* to the extension $\mathring{X}$ if there exist mutually inverse vector-lattice homomorphisms $v\colon \mathring{X} \to X$ and $w\colon X \to \mathring{X}$ such that X is larger than $\mathring{X}$ relative to v and $\mathring{X}$ is larger than X relative to w.

2.3. Let X be an extension of $C^*(T)$ inheriting separable decomposition.

The extension X will be called *saturated* if for any X_S and any proper component Y such that $Y^d \not\subset X_S$ there exists an X_R such that $X_S \cup Y \subset X_R$ and $R \subset S$, where $Y^d = \{x \in X | \forall y \in Y\ (|x| \wedge |y| = 0)\}$.

The extension X will be called *filled* if $\cap X_{S_k} = X_S$ for any sequence S_k such that $\cup S_k = S$.

The extension X will be called *σ-component* if every ideal X_S is a σ-component of X. The extension X will be called *lower component* if every ideal X_S is a lower component of X.

LEMMA. *Any saturated extension X is filled.*

PROOF. On the strength of Yosida's theorem ([12]) there is a compact H such that the vector-lattice X is isomorphic to the vector-lattice $C(H)$. Consider the non-empty closed subsets $H_S \equiv \{s \in H | \forall x \in X_S\ (x(s) = 0)\}$. Let $S = \cup S_k$. Then $\cup H_{S_k}$ is dense in H_S. In fact assume that there exists an open set G such that $G \cap (H_S \setminus \cup H_{S_k}) \neq \emptyset$. Take a regular closed set $F \subset G$ such that $H_S \cap \operatorname{int} F \neq \emptyset$. Consider the proper component $Y \equiv \{y \in X | y(F) = 0\}$. Then there exists an $R \subset S$ such that $X_S \cup Y \subset X_R$. So $H_R \subset G$. As $R_k \equiv R \cap S_k \neq \emptyset$ for some k we get $\emptyset \neq H_R \subset \subset H_R \cap H_{S_k} = \emptyset$. From this contradiction we conclude that such a set G does not exist.

Now take a $0 \leq x \in \cap X_{S_k}$. Then $x(H_S) = 0$. Consider the functions $x_k \equiv \equiv \left(x - \frac{1}{k}\mathbf{1}\right) \vee 0$. From the property $H_S \cap \operatorname{cl} \operatorname{coz} x_k = \emptyset$ we conclude that $x_k \in X_S$. As this ideal is uniformly closed we get $x \in X_S$. The lemma is proved.

3. Lattice-ring extension inheriting separable decomposition

We shall suppose that all considered f-rings are commutative, Archimedian, have strong units[1] and are uniformly complete with respect to their units and that all considered f-ring homomorphisms are unitary. Also we shall suppose that all considered f-ring ideals are uniformly closed.

[1] The unit $\mathbf{1}$ of an f-ring X is called the *strong unit*, if for any $x \in X$ there exists a natural number $n = n(x)$ such that $|x| \leq n\mathbf{1}$.

Let T be a completely regular space and $C^*(T)$ be the f-ring of all bounded continuous functions on T. Let X be an f-ring and $u\colon C^*(T)\to X$ be an injective f-ring homomorphism We shall say that X is an *f-ring extension* of $C^*(T)$ and shall identify $C^*(T)$ with its image in X.

3.1. If Y and Z are modules over the f-ring X then the set of all module homomorphisms from Y into Z is denoted by $\mathrm{Hom}_X(Y, Z)$. Let Y and Z be ring ideals in the f-ring X. A homomorphism $g\in\mathrm{Hom}_X(Y, Z)$ will be called *bounded* if there is a natural number n such that $|gy|\leqq n|y|$ for any $y\in Y$. The subset of $\mathrm{Hom}_X(Y, Z)$ consisting of all the bounded homomorphisms will be denoted by $\mathrm{Hom}_X^*(Y, Z)$.

The first and second annihilator of a subset Y of X will be denoted as usual by Y^* and Y^{**}, resp.

The f-ring X will be called *σ-continuing* if for any countably generated ring ideal Y of X and for any homomorphism $g\in\mathrm{Hom}_X^*(Y, Y^{**})$ there exists a homomorphism $h\in\mathrm{Hom}_X^*(X, Y^{**})$ extending g.

The extension X will be called *lower continuing* if for any ring ideal Y of the ring $C^*(T)$ and for any homomorphism $g\in\mathrm{Hom}_{C^*(T)}^*(Y, C^*(T)\cap Y^{**})$ there exists a homomorphism $h\in\mathrm{Hom}_X^*(X, Y^{**})$ extending g.

An f-ring ideal Z in X will be called a *σ-segment* of X if for any countably generated ring ideal Y of X and for any pair of homomorphisms $g\in\mathrm{Hom}_X^*(Y, Y^{**})$ and $h\in\mathrm{Hom}_X^*(X, Y^{**})$ such that h extends g the condition $g(Y)\subset Z$ implies $h(X)\subset Z$. The f-ring ideal Z will be called a *lower segment* of X if for any ring ideal of the ring $C^*(T)$ and for any pair of homomorphisms $g\in\mathrm{Hom}_{C^*(T)}^*(Y, C^*(T)\cap Y^{**})$ and $h\in\mathrm{Hom}_X^*(X, Y^{**})$ such that h extends g the condition $g(Y)\subset Z$ implies $h(X)\subset Z$.

3.2. For any countable set $S\in\mathscr{S}(T)$ consider the f-ring ideal $C_S^*(T)\equiv\equiv\{f\in C^*(T)|\, f(T_S)=0\}$ in the f-ring $C^*(T)$. The family $\{C_S^*(T)|S\in\mathscr{S}(T)\}$ will be called the *separable decomposition of the f-ring* $C^*(T)$.

The extension X of $C^*(T)$ will be called *inheriting separable decomposition* if X has a family of proper f-ring ideals $\{X_S|S\in\mathscr{S}(T)\}$ such that $\bigcap X_S=\{0\}$, $uf\in X_S$ iff $f\in C_S^*(T)$ and $S_1\subset S_2$ implies $X_{S_1}\supset X_{S_2}$. The mapping $C_S^*(T)\mapsto X_S$ will be called the *inheritance of separable decomposition*.

Let $\{X, u\colon C^*(T)\to X,\ C_S^*(T)\mapsto X_S\}$ and $\{\dot X, \dot u\colon C^*(T)\to\dot X,\ C_S^*(T)\mapsto\dot X_S\}$ be f-ring extensions inheriting separable decomposition. The extension X will be called *larger* than the extension $\dot X$ if there exists an injective f-ring homomorphism $v\colon\dot X\to X$ such that $v\circ\dot u=u$ and $v(\dot X_S)\subset X_S$. The extension X will be called *isomorphic* to the extension $\dot X$ if there exist mutually inverse f-ring homomorphisms $v\colon\dot X\to X$ and $w\colon X\to\dot X$ such that X is larger than $\dot X$ relative to v and $\dot X$ is larger than X relative to w.

3.3. Let X be an f-ring extension of $C^*(T)$ inheriting separable decomposition.

An f-ring ideal Y in X is called an *annihilator f-ring ideal* if Y coincides with its own second annihilator Y^{**}.

The extension X will be called *saturated* if for any X_S and any proper annihilator f-ring ideal Y such that $Y^*\not\subset X_S$ there exists an X_R such that $X_S\cup Y\subset X_R$ and $R\subset S$.

The extension X will be called *filled* if $\bigcap X_{S_k}=X_S$ for any sequence S_k such that $\bigcup S_k=S$. Any saturated extension is filled.

The extension X will be called *σ-segment* if any X_S is a σ-segment of X. The extension X will be called *lower segment* if any X_S is a lower segment of X.

§ 1. Borelian cover

Let T be a completely regular space. The set of all the Borel subsets of T will be denoted by $\mathscr{B}(T)$.

Consider the Stone compact H_0 of all ultrafilters in $\mathscr{B}(T)$. For any point $s\in H_0$ let P_s denote the set $\bigcap\{\mathrm{cl}\,B\,|\,B\in\Theta_s\}$ where s corresponds to the ultrafilter Θ_s. Consider the subspace $H\equiv\{s\in H_0|P_s\neq\emptyset\}$ and define the surjective continuous mapping $\eta\colon H\to T$ such that $\eta s\equiv P_s$. The space H with the mapping η will be called the *Borelian cover* of T.

Let i_0 be the Stone isomorphism between $\mathscr{B}(T)$ and the Boolean algebra $\Delta(H_0)$ of all open-closed subsets of H_0. Let $i\colon \mathscr{B}(T)\to\Delta(H)$ be the corresponding homomorphism of Boolean algebras such that $iB\equiv H\cap i_0B$.

It can be checked that the subspace H is dense in H_0, the homomorphism i is injective and the mapping η is perfect.

Associate with a countable set S the closed subspace H_S of all the ultrafilters from H, all the members of which intersect the set S. Then the set $H_t=it$ is an isolated point in H for every point $t\in T$, $\bigcup\{H_t|t\in T\}$ is dense in H and there are no other isolated points in H.

Borel sets B_1 and B_2 will be called *S-equivalent* if $(B\triangle B')\cap S=\emptyset$. Let $\mathscr{B}_S(T)$ denote the Boolean algebra of all classes of S-equivalence $\bar{B}$ of elements B from $\mathscr{B}(T)$. Let $(B\triangle B')\cap S=\emptyset$. Let $s\in iB\cap H_S$ and s correspond to an ultrafilter Θ_s. Assume $B'\notin\Theta_s$. Then $B\triangle B'\supset B\setminus B'\in\Theta_s$ but this is false. Hence $s\in iB'\cap H_S$. Thus $\mathrm{i}B\cap H_S=\mathrm{i}B'\cap H_S$. So we can define correctly the homomorphism of Boolean algebras $\mathrm{i}_S\colon \mathscr{B}_S(T)\to\Delta(H_S)$ such that $\mathrm{i}_S\bar{B}\equiv\mathrm{i}B\cap H_S$.

LEMMA 1. *The space H_S is extremally disconnected, i_S is an isomorphism and $\eta H_S=T_S$.*

PROOF. Let Q_S denote the Stone compact of all the ultrafilters of the Boolean algebra $\mathscr{B}_S(T)$. As this Boolean algebra is complete the space Q_S is extremally disconnected. Denote by H_{0_S} the subspace of H_0 consisting of all the ultrafilters, all the members of which intersect the set S. Consider the Boolean homomorphism $h_S\colon \mathscr{B}(T)\to\mathscr{B}_S(T)$ such that $h_SB\equiv\bar{B}$. Let Θ' be an ultrafilter in $\mathscr{B}_S(T)$. Then $\Theta\equiv h_S^{-1}\Theta'$ is an ultrafilter in $\mathscr{B}(T)$, $\Theta\in H_{0_S}$ and the mapping $\gamma_S\colon \Theta'\to\Theta$ is an injective continuous mapping from Q_S onto H_{0_S}. This implies that the space H_{0_S} is extremally disconnected.

The homomorphism i_S is injective. In fact let $\mathrm{i}_S\bar{B}=\emptyset$ and assume $B\cap S\neq\emptyset$. Consider a compact set $F\subset B\cap S$. Consider the proper filter base Θ_0 in $\mathscr{B}(T)$ consisting of the set F and all open sets G containing F. Then $\Theta_0'\equiv h_S\Theta_0$ is a proper filter base in $\mathscr{B}_S(T)$. Imbed Θ_0' in some ultrafilter Θ' and consider the ultrafilter $\Theta\equiv h_S^{-1}\Theta'\in H_{0_S}$. If $B'\in\Theta$ then $F\cap\mathrm{cl}\,B'\neq\emptyset$. Otherwise $G\cap\mathrm{cl}\,B'=\emptyset$ for some $G\in\Theta_0\subset\Theta$. This implies $B'\cap G\in\Theta$. Hence $\emptyset=B'\cap G\cap S\neq\emptyset$. It follows from this contradiction that $\bigcap\{F\cap \text{cl } B'|B'\in\Theta\}\neq\emptyset$. Therefore $\Theta\in H_S$. Besides this $\bar{F}\in\Theta'$ implies $\Theta\in \mathrm{i}_S\bar{B}=\emptyset$. Thus $\bar{B}=0$.

Define the homomorphism $\mathrm{i}_{0_S}: \mathscr{B}_S(T)\to\Delta(H_{0_S})$ by setting $\mathrm{i}_{0_S}\bar{B}\equiv \mathrm{i}_0 B\cap H_{0_S}$. Let U be an arbitrary open-closed subset of H_{0_S}. Consider $V\equiv\gamma_S^{-1}U$. Then $V=\{\Theta'\in Q_S|\bar{B}\in\Theta'\}$ for some $\bar{B}\neq 0$. Take $\Theta\in \mathrm{i}_{0_S}\bar{B}$ and $\Theta'\equiv h_S\Theta\in V$. Since $\Theta=h_S^{-1}\Theta'$ we have $\Theta\in U$. On the other hand let $\Theta\in U$ and $\Theta'\equiv\gamma_S^{-1}\Theta$. Then $\bar{B}\in\Theta'$ implies $\Theta\in \mathrm{i}_{0_S}\bar{B}$. Thus $U=\mathrm{i}_{0_S}\bar{B}$. So $U\cap H_S=\mathrm{i}_S\bar{B}\neq\emptyset$. It means that H_S is dense in H_{0_S}. Therefore the space H_S is extremally disconnected.

Let U be an open-closed subset of the space H_S, V be its complement, U' and V' be the closures of U and V in the space H_{0_S}. As $H_{0_S}=\beta H_S$ we have $U'\cap V'=\emptyset$ and $H_{0_S}=U'\cup V'$. So U' is open-closed. It follows from above that $U'=\mathrm{i}_{0_S}\bar{B}$. Hence $U=\mathrm{i}_S\bar{B}$. Thus the homomorphism i_S is surjective.

Let $t\in T_S$. Consider the proper filter base Θ_0 consisting of all open sets G containing t. Imbed $h_S\Theta_0$ in some ultrafilter Θ' and consider the ultrafilter $\Theta\equiv h_S^{-1}\Theta'\in H_{0_S}$. As $t\in\bigcap\{\text{cl } B|B\in\Theta\}$ we have $\Theta\in H_S$ and $\eta\Theta=t$. Further let a $\Theta\in H_S$ and G be a neighbourhood of $\eta\Theta$. As $G\cap B\neq\emptyset$ for any $B\in\Theta$ we obtain $G\in\Theta$. So $G\cap T_S\neq\emptyset$. It means $\eta\Theta\in T_S$. Thus $\eta H_S=T_S$. The lemma is proved.

COROLLARY. *The space H_S satisfies the Souslin condition, i.e. each family of disjoint open subsets in H_S is countable.*

LEMMA 2. *In H_S any meager subset is nowhere dense.*

PROOF. Let F_k be a sequence of closed nowhere dense subsets in H_S. Assume that $\text{cl}\bigcup F_k$ is not nowhere dense, i.e. there is an open-closed set $U\equiv \mathrm{i}_S\bar{B}$ in H_S such that $U\subset\text{cl}\bigcup F_k$. We can suppose that $F_k\subset U$ and so $U=\text{cl}\bigcup F_k$. According to the previous corollary for every k there exists in U a sequence $V_{kj}\equiv \mathrm{i}_S\bar{B}_{kj}$ of decreasing open-closed subsets with nowhere dense intersection, containing the set F_k. We can suppose that $B\supset B_{kj}\supset B_{k,j+1}$. Assume that $E_k\equiv\bigcap B_{kj}\cap S\neq\emptyset$. Then $\emptyset\neq \mathrm{i}_S\bar{E}_k\subset V_{kj}$ for any j, but this is false. So $E_k=\emptyset$. Take a point $t\in S$. Then there exists a set B_{kj_k} such that t does not belong to this set. Consider the non-empty set $V\equiv \mathrm{i}_S\bar{t}$. We get $\bigcup F_k\cap V\subset\bigcup(V_{kj_k}\cap V)=\emptyset$ and $V\subset U$, but this is impossible.

COROLLARY. *The space H_S is Baire.*

Now we need a classification of Borel sets. The classification of Young ([6]), used usually in mathematical literature, is not suitable for us since it is valid only for such spaces whose open sets have the type F_σ. It can be checked that the following classification is valid. $\mathscr{B}(T)=\bigcup\{\mathscr{B}_\alpha(T)|\alpha<\omega_1\}$, where $\mathscr{B}_0(T)$ consists of all open subsets of T and

$$\mathscr{B}_\alpha(T)\equiv\Big\{\bigcup_k(B_k\cup(T\setminus C_k))|\exists\beta_k<\alpha\ \exists\gamma_k<\alpha\ \ (B_k\in\mathscr{B}_{\beta_k}(T), C_k\in\mathscr{B}_{\gamma_k}(T))\Big\}.$$

Denote by η_S the restriction of η to H_S.

LEMMA 3. *For any Borel set B the set $\mathrm{i}_S\bar{B}\triangle\eta_S^{-1}B$ is nowhere dense in the space H_S.*

PROOF. Let G be an open set. Then $\mathrm{i}_S\bar{G}=\text{cl }\eta_S^{-1}G$. In fact assume that there exists $s\in\eta_S^{-1}G\cap \mathrm{i}_S(\overline{T\setminus G})$. Then $\eta_S s\in T\setminus G$ but this is false. On the other hand assume that there exists a non-empty set $U \equiv \mathrm{i}_S \bar{B}$ such that $U \subset \mathrm{i}_S \bar{G} \setminus \mathrm{cl}\, \eta_S^{-1} G$ and $B \subset G$. Represent G in the form $G = \bigcup F_p$ for some family of closed sets F_p. As $B \cap F_p \cap S \neq \emptyset$ for some p we can consider the non-empty set $V \equiv \mathrm{i}_S(\overline{B \cap F_p})$. Then $\emptyset \neq \eta_S V \subset F_p \subset G$. On the other hand $V \subset U \subset \eta_S^{-1}(T \setminus G)$, but this is impossible. Thus the assertion of the lemma is valid for the 0-th Borel class. It is proved by induction that the same is valid for all other classes.

Corollary 1. *Let B be a Borel set. Then $B \cap S = \emptyset$ iff $\eta_S^{-1} B$ is nowhere dense in H_S.*

Corollary 2. *Let B be a Borel set. Then $\mathrm{i}B \triangle \eta^{-1} B \not\supset H_S$ for any S.*

Lemma 4. *For any H_S and any open-closed set U in H there exists a countable set $R \subset S$ such that $H_S \cap U = H_R$.*

Proof. Consider the non-empty set $V \equiv H_S \cap U$. By Lemma 1 $V = \mathrm{i}_S \bar{B}$ for some B. Consider the set $R \equiv B \cap S$. Then $H_R \subset H_S$. Let $\Theta \in V$. Then $B \in \Theta$. So $B' \cap R \neq \emptyset$ for any $B' \in \Theta$. This means $\Theta \in H_R$. Conversely, let $\Theta \in H_R$. For any $B' \in \Theta$ we have $B' \cap B \cap S \neq \emptyset$. It has a consequence that $B \in \Theta$ and so $\Theta \in V$.

Lemma 5. *If $S_1 \subset S_2$ then H_{S_1} is an open-closed subset in the space H_{S_2}.*

Proof. Consider the set $S \equiv S_2 \setminus S_1$. Then $H_S \cup H_{S_1} \subset H_{S_2}$. Assume that there exists an ultrafilter $\Theta \in H_S \cap H_{S_1}$. For any point $t \in S_1$ consider an open set G_t such that $t \notin G_t$ and $S \subset G_t$. Let $B \equiv \bigcap \{G_t | t \in S_1\}$. Then $B \cap S_1 = \emptyset$ and $(T \setminus B) \cap \cap S = \emptyset$ imply that $B \notin \Theta$ and $T \setminus B \notin \Theta$ but this is impossible. So $H_S \cap H_{S_1} = \emptyset$. Assume $E \equiv H_{S_2} \setminus (H_S \cup H_{S_1}) \neq \emptyset$. Then by the previous lemma there exists a subset $R \subset S_2$ such that $H_R \subset E$. Denote $R \cap S_1 \equiv R_1$ and $R \cap S \equiv R_2$. Then $\emptyset \neq H_{R_1} \cup \cup H_{R_2} \subset H_R \cap (H_S \cup H_{S_2}) = \emptyset$. It follows from this contradiction that $E = \emptyset$. The lemma is proved.

Consider in the space H the base $\mathscr{C}_0(H)$ consisting of all cozero-sets C which can be represented in the form $C = \bigcup U_k$ for some sequence of open-closed subsets $U_n \in \mathrm{i}\mathscr{B}(T)$. Then $\Delta_0(H) = \mathrm{i}\mathscr{B}(T)$.

Lemma 6. *The base $\mathscr{C}_0(H)$ is completely normal.*

Proof. Consider the uniformly complete vector-lattice

$$\Phi \equiv \{f \in C^*(H) | \exists f' \in C^*(H_0)(f = f'|H)\}.$$

Let $C = \bigcup \mathrm{i}B_k$. Consider the functions $f \equiv \sum \chi(\mathrm{i}B_k)/2^k \in C^*(H)$ and

$$f' \equiv \sum \chi(\mathrm{i}_0 B_k)/2^k \in C^*(H_0).$$

As $f = f'|H$ we have $f \in \Phi$. In addition $C = \mathrm{coz} f$. Conversely let $C = \mathrm{coz} f$ for some $f \in \Phi$. Then $f = f'|H$ for some $f' \in C^*(H_0)$. As $\mathrm{coz} f' = \bigcup \mathrm{i}_0 B_k$ for some sequence B_k, we get $C = \bigcup \mathrm{i}B_k$. Thus $\mathscr{C}_0(H) = \{\mathrm{coz} f | f \in \Phi\}$. This has as a consequence that the base is completely normal.

Lemma 7. *Let G be an open subset of T. Then $\mathrm{cl}\, \eta^{-1} G = \mathrm{i}G$.*

Proof. Assume that there exists a point $s \in \eta^{-1} G \cap \mathrm{i}(T \setminus G)$. Then $\eta s \in G \cap \cap (T \setminus G)$ but this is impossible. On the other hand assume that there exists a non-empty set $U \equiv \mathrm{i}B$ such that $U \subset \mathrm{i}G \setminus \mathrm{cl}\,\eta^{-1}G$. We can suppose that $B \subset G$. Represent G in the form $G = \bigcup F_p$ for some family of closed sets F_p. Then $V \equiv \mathrm{i}(B \cap F_p) \neq \emptyset$ for some p. This implies $\emptyset \neq \eta V \subset F_p \subset G$. But this contradicts to the inclusion $V \subset \eta^{-1}(T \setminus G)$.

COROLLARY 1. *The preimage H is lower extremally disconnected.*

COROLLARY 2. *The preimage H is lower disjoined.*

PROOF. Let $\eta^{-1}G \cap H_S = \emptyset$. By the corollary to Lemma 3 $G \cap S = \emptyset$. This implies $\mathrm{i}G \cap H_S = \emptyset$.

LEMMA 8. *The preimage H is σ-extremally disconnected and σ-disjoined.*

PROOF. Let $C = \bigcup \mathrm{i}B_k$. Consider the sets $B \equiv \bigcup B_k$ and $U \equiv \mathrm{i}B$. It is evident that $\mathrm{cl}\,C = U$.

Let $C \cap H_S = \emptyset$. Then according to Lemma 1 $B_k \cap S = \emptyset$. So $B \cap S = \emptyset$ implies $U \cap H_S = \emptyset$ in the same manner.

Now let H be an arbitrary preimage of T lifting separable covering and having a completely normal base $\mathscr{C}_0(H)$. The preimage H will be called *Borel determined* if for any cozero-set $C \in \mathscr{C}_0(H)$ there exists a $B \in \mathscr{B}(T)$ such that $C \triangle \eta^{-1}B \not\supset H_S$ for any $S \in \mathscr{S}(T)$.

LEMMA 9. *Let H be the Borelian cover of T. Then the preimage H is Borel determined.*

PROOF. Let $C = \bigcup \mathrm{i}B_k$ and $B \equiv \bigcup B_k$. Then according to Lemma 3 and 2 the set $(C \triangle \eta^{-1}B) \cap H_S$ is nowhere-dense in H_S. So $C \triangle \eta^{-1}B \not\supset H_S$.

Further uniqueness is understood up to isomorphism.

THEOREM 1. *Let H be the Borelian cover of T. Then*

(1) *H is the unique largest of all the perfect saturated Borel determined preimages of T lifting separable covering;*

(2) *H is the unique smallest of all the perfect filled lower extremally disconnected σ-extremally disconnected lower disjoined σ-disjoined preimages of T lifting separable covering and moreover H is the unique universal (in the sense of Bourbaki) among all such preimages;*

(3) *H is the unique perfect saturated Borel determined lower extremally disconnected σ-extremally disconnected lower disjoined σ-disjoined preimage of T lifting separable covering.*

PROOF. Let $\{\dot{H}, \dot{\eta}\colon \dot{H} \to T,\ T_S \mapsto \dot{H}_S, \mathscr{C}_0(\dot{H})\}$ be a preimage of T having the properties from 1). Then for any C from the base there is a Borel set B such that $C \triangle \dot{\eta}^{-1}B \not\supset \dot{H}_S$ for any S. Assume that there exists another set B_1 having this property. Assume that there exists a point $t \in B \setminus B_1$. Then $\dot{H}_t \subset \dot{\eta}^{-1}B \cap (\dot{H} \setminus \dot{\eta}^{-1}B_1)$. By the given property $\dot{H}_t \cap C \neq \emptyset$. Therefore $\dot{H}_t \subset C \setminus \dot{\eta}^{-1}B_1$ but this is false. That is why $B = B_1$. So we can define correctly the mapping $k\colon \mathscr{C}_0(\dot{H}) \to \mathscr{B}(T)$ such that $kC \equiv B$.

Verify that k is a lattice homomorphism. Let $kC_1 = B_1$ and $kC_2 = B_2$. Assume that $\dot{H}_S \subset (C_1 \cup C_2) \triangle \dot{\eta}^{-1}(B_1 \cup B_2)$. If $\dot{H}_S \cap ((C_1 \cup C_2) \setminus \dot{\eta}^{-1}(B_1 \cup B_2)) \neq \emptyset$ then there exists a set $R\subset S$ such that $\dot{H}_R\subset\dot{H}_S\cap(C_1\cup C_2)$. Let $\dot{H}_R\cap C_1\neq\emptyset$. Then there exists a set $R_1\subset R$ such that $\dot{H}_{R_1}\subset\dot{H}_R\cap C_1\subset C_1\setminus\dot{\eta}^{-1}B_1$ but this is impossible. Hence $\dot{H}_S\subset\dot{\eta}^{-1}(B_1\cup B_2)\setminus(C_1\cup C_2)$. Then the inclusion $T_S\subset B_1\cup B_2$ means that $B_1\cap S\neq\emptyset$ for example. Take a compact subset R from this intersection. Then $\dot{H}_R\subset\dot{\eta}^{-1}B_1\cap\dot{H}_S\subset\dot{\eta}^{-1}B_1\setminus C_1$ but this is false. As a result we get $k(C_1\cup C_2)=$ $=B_1\cup B_2$.

Assume that $\dot{H}_S\subset(C_1\cap C_2)\triangle\dot{\eta}^{-1}(B_1\cap B_2)$. If $\dot{H}_S\cap((C_1\cap C_2)\setminus\dot{\eta}^{-1}(B_1\cap B_2))\neq\emptyset$ then there exists a set $R\subset S$ such that $\dot{H}_R\subset\dot{H}_S\cap(C_1\cap C_2)$. This implies $\dot{H}_R\cap$ $\cap\dot{\eta}^{-1}B_1\neq\emptyset$. Assume $B_1\cap R=\emptyset$. Then there exists a compact set $R_1\subset(T\setminus B_1)\cap R$. So $\dot{H}_{R_1}\subset\dot{H}_R\cap(\dot{H}\setminus\dot{\eta}^{-1}B_1)\subset C_1\setminus\dot{\eta}^{-1}B_1$ but this is false. Hence $B_1\cap R\neq\emptyset$. But then similarly $\emptyset\neq T_{R_2}\subset B_1$ for some $R_2\subset R$. So $\dot{H}_{R_2}\subset\dot{H}_R\cap\dot{\eta}^{-1}B_1$. As $\dot{H}_R\cap\dot{\eta}^{-1}(B_1\cap B_2)\neq\emptyset$ we get $\dot{H}_{R_2}\subset C_2\setminus\dot{\eta}^{-1}B_2$. It follows from this contradiction that

$$\dot{H}_S \subset \dot{\eta}^{-1}(B_1\cap B_2)\setminus(C_1\cap C_2).$$

Then $\dot{H}_S\cap C_1\neq\emptyset$. So there exists a set $R\subset S$ such that $\dot{H}_R\subset\dot{H}_S\cap C_1$. But then $\dot{H}_R\subset\dot{\eta}^{-1}B_2\setminus C_2$. As this is impossible we get as a result that $k(C_1\cap C_2)=B_1\cap B_2$. It is clear that k preserves the unit.

Let $C\neq\emptyset$. Then $C\cap\dot{H}_S\neq\emptyset$ for some set S. So there exists a set $R\subset S$ such that $\dot{H}_R\subset\dot{H}_S\cap C$. This implies $kC\neq\emptyset$. Conversely, let $kC\neq\emptyset$. Take a point $t\in kC$. Then $\dot{H}_t\subset\dot{\eta}^{-1}kC$ implies $C\neq\emptyset$.

Check that $\mathrm{cl}\,\dot{\eta}C=\mathrm{cl}\,kC$. Denote the left-hand set by P and the right-hand set by Q. Let $s\in C$ and $t\equiv\dot{\eta}s\notin Q$. Then there exists a cozero-set G such that $t\in G\subset$ $\subset T\setminus Q$. Denote $C_1\equiv\dot{\eta}^{-1}G$. We have $\emptyset\neq k(C_1\cap C)=kC\cap G=\emptyset$. From this contradiction we get $\dot{\eta}C\subset Q$. Now assume that there exists a cozero-set G such that $G\cap P=\emptyset$ and $G\cap Q\neq\emptyset$. Denote $C_1\equiv\dot{\eta}^{-1}G$. We get $C\cap C_1\neq\emptyset$ because of $k(C_1\cap C)=G\cap kC\neq\emptyset$. But this is false.

Now let $\{H,\eta\colon H\to T,\ T_S\mapsto H_S,\ \mathscr{C}_0(H)\}$ be a preimage of T with the properties from 2).

Assume that for some Borel set B there exists a set $U\in\Delta_0(H)$ such that for any S there exists a representation $S=(\cup S_j)\cup(\cup S_k)$ for some sequences of compact sets S_j and S_k such that $\cup H_j\subset U\cap\eta^{-1}B\subset U\cup\eta^{-1}B\subset H\setminus\cup H_k$ where $H_r\equiv$ $\equiv H_{S_r}$ for $r\in\{j,k\}$. Then the set U with this property is unique. In fact let for B there exists another set $U'\in\Delta_0(H)$ such that for any S there exist a representation $S=\cup S_p\cup\cup S_q$ such that $\cup H_p\subset U'\cap\eta^{-1}B\subset U'\cup\eta^{-1}B\subset H\setminus\cup H_q$ where $H_r\equiv$ $\equiv H_{S_r}$ for $r\in\{p,q\}$. Consider the sets $S_{jp}\equiv S_j\cap S_p$. In the same manner we define the sets S_{jq}, S_{kp} and S_{kq}. Denote $H_{S_{rs}}$ by H_{rs} for any $r,s\in\{j,k,p,q\}$. Then $H_{rs}\subset$ $\subset H_r\cap H_s$. So $(\cup H_{jp})\cup(\cup H_{jq})\subset U\cap\eta^{-1}B\subset U\cup\eta^{-1}B\subset H\setminus(\cup H_{kp})\cup(\cup H_{kq})$ and similarly $(\cup H_{jp})\cup(\cup H_{kp})\subset U'\cap\eta^{-1}B\subset U'\cup\eta^{-1}B\subset H\setminus(\cup H_{jq})\cup(\cup H_{kq})$. So $U\triangle$ $\triangle U'\subset(U\triangle\eta^{-1}B)\cup(U'\triangle\eta^{-1}B)\subset H\setminus((\cup H_{jp})\cup(\cup H_{jq})\cup(\cup H_{kp})\cup(\cup H_{kq}))$. As the preimage H is filled we conclude that $(U\triangle U')\cap H_S=\emptyset$. Since this condition is fulfilled for any S we get $U=U'$.

Denote U by $\mathrm{i}B$. Prove that $\mathrm{i}B$ is defined for all Borel sets B. Let G be an open set. Then $G\cap S=\cup S_j$ and $(T\setminus G)\cap S=\cup S_k$ for some sequences of compact subsets of S. Consider the open-closed set $U\equiv\mathrm{cl}\,\eta^{-1}G\in\Delta_0(H)$. By virtue of the lower disjoinedness of the preimage H we get $\cup H_j\subset U\cap\eta^{-1}G\subset U\cup\eta^{-1}G\subset$ $\subset H\setminus\cup H_k$. Therefore $U=\mathrm{i}G$.

Further we shall proceed by induction. Assume that the set $\mathrm{i}B$ exists for any

$B\in\mathscr{B}_\beta(T)$ and for any $\beta<\alpha$. Let $B\in\mathscr{B}_\alpha(T)$, $B=\cup(B'_m\cup(T\setminus B''_m))$, $U'_m\equiv \mathrm{i}B'_m$ and $U''_m\equiv\mathrm{i}B''_m$. Then $H\setminus U''_m=\mathrm{i}(T\setminus B''_m)$. Denote $B_{2m}\equiv B'_m$, $B_{2m+1}\equiv T\setminus B''_m$, $U_{2m}\equiv U'_m$ and $U_{2m+1}\equiv H\setminus U''_m$. Then $U_m=\mathrm{i}B_m$. This means that $\cup H_j^m\subset U_m\cap\eta^{-1}B_m\subset U_m\cup\cup\eta^{-1}B_m\subset H\setminus\cup H_k^m$ for some representations $S=\cup S_j^m\cup\cup S_k^m$. Applying the mapping η to this chain of the inclusions we get $\cup S_j^m\subset B_m\subset T\setminus\cup S_k^m$. So $\cup S_j^m==B_m\cap S$. Represent $S\setminus B$ in the form $S\setminus B=\cup S_r$ for some sequence of compact sets S_r. Consider the sets $S_{kr}^m\equiv S_k^m\cap S_r$. As $\bigcup\limits_k S_{kr}^m=S_r$ the set $\bigcup\limits_k H_{kr}^m$ is dense in H_r. Since $U_m\cap\bigcup\limits_k H_{kr}^m\subset U_m\cap\bigcup\limits_k H_k^m=\emptyset$ we get $U\cap H_r=\emptyset$. Consider the open-closed set $U\equiv\mathrm{cl}\cup U_m\in\Delta_0(H)$. By virtue of the σ-disjoinedness of the preimage H we have $U\cap H_r=\emptyset$. Besides $\eta^{-1}B\cap H_r=\emptyset$. As a result we get $\bigcup\limits_m\bigcup\limits_j H_j^m\subset U\cap\cap\eta^{-1}B\subset U\cup\eta^{-1}B\subset H\setminus\cup H_r$. In addition for the set S we have the representation $S=(\bigcup\limits_m\bigcup\limits_j S_j^m)\cup(\cup S_r)$. This means that $U=\mathrm{i}B$.

Consequently we can consider the mapping $\mathrm{i}\colon\mathscr{B}(T)\to\Delta_0(H)$. Check that this mapping is a homomorphism of Boolean algebras. Let $\mathrm{i}B=U$ and $\mathrm{i}B'=U'$, i.e. $\cup H_j\subset U\cap\eta^{-1}B\subset U\cup\eta^{-1}B\subset H\setminus\cup H_k$ for some representation $S=\cup S_j\cup\cup S_k$ and $\cup H_p\subset U'\cap\eta^{-1}B\subset U'\cap\eta^{-1}B'\subset H\setminus\cup H_q$ for some representation $S=\cup S_p\cup\cup S_q$. Consider as it was done above the sets S_{rs} and H_{rs} for any $r,s\in\{j,k,p,q\}$. Then

$$(\cup H_{jp})\cup(\cup H_{jq})\cup(\cup H_{kp})\subset(\cup H_j)\cup(\cup H_p)\subset(U\cup U')\cap\eta^{-1}(B\cup B')\subset$$
$$\subset(U\cup U')\cup\eta^{-1}(B\cup B')\subset H\setminus((\cup H_k)\cap(\cup H_q))\subset H\setminus\cup H_{kq}$$

means that $\mathrm{i}(B\cup B')=U\cup U'$. Hence i preserves the supremum. Let $U=\mathrm{i}T$. Then $\cup H_j\subset U\subset H\subset H\setminus\cup H_k$ implies that $U\cap H_S=H_S$ for any S. So $U=H$. Hence i preserves the unit. It is obvious that i preserves the complement.

If $B\neq\emptyset$ then for some point $t\in B$ we consider the set $S\equiv\{t\}$. In this case $S_j\neq\emptyset$ implies $H_j\neq\emptyset$. So $U\neq\emptyset$. Thus i is injective.

Consider the unit preserving lattice homomorphism $\alpha\colon\mathscr{C}_0(\dot H)\to\Delta_0(H)$ such that $\alpha\equiv\mathrm{i}\circ k$. Let $t\in H$. Consider the sets $\Gamma\equiv\{C\in\mathscr{C}_0(\dot H)|t\in\alpha C\}$ and $P\equiv\dot\eta^{-1}\eta t$. Assume that $P\cap\mathrm{cl}\,C=\emptyset$ for some $C\in\Gamma$. Then $\eta t\in G\equiv T\setminus\mathrm{cl}\,\dot\eta C$ implies $t\in\eta^{-1}G\subset\subset\mathrm{i}G$. As $G\cap\mathrm{cl}\,kC=\emptyset$ we have $\mathrm{i}G\cap\alpha C=\emptyset$ but this is false. Hence $P\cap\mathrm{cl}\,C\neq\emptyset$ for any $C\in\Gamma$. Let $C_1,C_2\in\Gamma$. Then $C_1\cap C_2\in\Gamma$ implies $P\cap\mathrm{cl}\,C_1\cap\mathrm{cl}\,C_2\neq\emptyset$. Therefore $P_t\equiv\cap\{P\cap\mathrm{cl}\,C|C\in\Gamma\}\neq\emptyset$ because of the compactness of P. Assume that there exist $s_1,s_2\in P_t$. As the base $\mathscr{C}_0(\dot K)$ is completely normal by the lemma from 1.1 there exist cozero-sets C_1 and C_2 from this base such that $s_1\in C_1$, $s_2\in C_2$, $s_1\notin\mathrm{cl}\,C_2$, $s_2\notin\mathrm{cl}\,C_1$ and $C_1\cup C_2=\dot H$. Then $H=\alpha\dot H=\alpha C_1\cup\alpha C_2$. Assume that $t\in\alpha C_1$. Then $s_2\in\mathrm{cl}\,C_1$ but this is false. It means that the set P_t consists of only one point. So we can define correctly the mapping $\gamma\colon H\to\dot H$ by means of the equality $\gamma t\equiv P_t$. This mapping is continuous. In fact let G be a neighbourhood of the point $s\equiv\gamma t$. Consider a cozero-set C from our base such that $s\in C\subset\mathrm{cl}\,C\subset G$. We can take a cozero-set C_1 from the base such that $s\notin\mathrm{cl}\,C_1$ and $C\cup C_1=\dot H$ according to the lemma from 1.1. This implies $H=\alpha C\cup\alpha C_1$. The assumption $t\in\alpha C_1$ implies $s\in\mathrm{cl}\,C_1$ but this is false. Hence $t\in\alpha C$. Let $t_1\in\alpha C$. Then $\gamma t_1\in\mathrm{cl}\,C\subset G$. As the set αC is open we obtain the continuity of γ.

This mapping is surjective. In fact consider a point $s\in\dot H$, the set

$$\Gamma\equiv\{C\in\mathscr{C}_0(\dot H)|s\in C\}$$

and the set $P \equiv \eta^{-1}\dot{\eta}s$. Assume that $\alpha C \cap P = \emptyset$ for some $C \in \Gamma$. Then there exists a cozero-set G such that $\dot{\eta}s \in G \subset T \setminus \eta\alpha C$. Then $s \in \dot{\eta}^{-1}G \equiv C_1$. On the other hand $\eta^{-1}G \cap \alpha C = \emptyset$ implies $\alpha(C_1 \cap C) = \mathrm{i}kC_1 \cap \alpha C = \mathrm{i}G \cap \alpha C = \emptyset$. It has as a consequence that $C_1 \cap C = \emptyset$ but this is false. It follows from this contradiction that $\alpha C \cap P \neq \emptyset$ for any $C \in \Gamma$. Therefore there exists a point $t \in \cap \{\alpha C \cap P | C \in \Gamma\}$. Consequently, $\gamma t \in \cap \{\mathrm{cl}\, C | C \in \Gamma\} = s$.

From the definition of the mapping γ we conclude that $\dot{\eta} \circ \gamma = \eta$. It has as a consequence that this mapping is perfect ([7], VI, § 2, 56).

Prove that $\gamma H_S = \dot{H}_S$. Assume that there exists a cozero-set $C \in \mathscr{C}_0(\dot{H})$ such that $C \cap \dot{H}_S \neq \emptyset$ and $\mathrm{cl}\, C \cap \gamma H_S = \emptyset$. Assume that there exists a point $t \in \alpha C \cap H_S$. Then $\gamma t \in \mathrm{cl}\, C \cap \gamma H_S = \emptyset$ but this is impossible. Therefore $\alpha C \cap H_S = \emptyset$. Consider the set $B \equiv kC$. Consider for B the corresponding sets S_j and $H_j \subset H_S$ defined above. Then $\cup H_j \subset \mathrm{i}B$ implies $H_j = \emptyset$. So $S_j = \emptyset$. Therefore $B \cap S = \emptyset$. On the other hand there exists a set $R \subset S$ such that $\dot{H}_R \subset C \cap \dot{H}_S$. Take a compact set $R_1 \subset R$. As $B \cap R_1 = \emptyset$ we have $\dot{H}_{R_1} \subset (\dot{H} \setminus \dot{\eta}^{-1}B) \cap \dot{H}_R \subset C \setminus \dot{\eta}^{-1}B$ but this is impossible. From this contradiction we conclude that $\dot{H}_S \subset \gamma H_S$.

Conversely assume that there exists a cozero-set $C \in \mathscr{C}_0(\dot{H})$ such that $C \cap \gamma H_S \neq \emptyset$ and $C \cap \dot{H}_S = \emptyset$. Consider the set $B \equiv kC$. Assume that $B \cap S \neq \emptyset$. Then there exists a compact set $R \subset B \cap S$. This implies $\dot{H}_R \subset \dot{\eta}^{-1}B \cap \dot{H}_S \subset \dot{\eta}^{-1}B \setminus C$ but this is impossible. So $B \cap S = \emptyset$. Consider for B the corresponding sets S_j and S_k defined above. As $S_j = \emptyset$ and $S = \cup S_k$ we have $\mathrm{i}B \subset H \setminus \cup H_k$. This implies $\mathrm{i}B \cap H_S = \emptyset$. Let t be a point of $\gamma^{-1}C$. Then there exists a cozero-set C_1 from the base such that $\gamma t \notin \mathrm{cl}\, C_1$ and $C \cup C_1 = \dot{H}$. Then $\alpha C \cup \alpha C_1 = H$ shows that $t \in \alpha C$. That is why $\gamma^{-1}C \subset \alpha C$. So we get $C \cap \gamma H_S = \emptyset$ but this contradicts to our assumption.

Thus H is larger than $\dot{H}$. Now let H be the Borelian cover of T. As the Borelian cover has the properties from 1) and 2) simultaneously we get as a result that the Borelian cover is the largest of all the preimages with the properties from 1) and the smallest of all the preimages with the properties from 2).

Let $\dot{H}$ be some other largest preimage of T. Then there are mappings $\gamma\colon H \to \dot{H}$ and $\delta\colon \dot{H} \to H$ such that H is larger than $\dot{H}$ relative to γ and $\dot{H}$ is larger than H relative to δ. Let $t \in H_S$. Then $t = \cap \{H_R\}$. This implies $\gamma t \in \cap \{\dot{H}_R\}$ and $\delta\gamma t \in \cap \cap \{H_R\} = t$. As $\cup H_S$ is dense we conclude that $\delta \circ \gamma = \mathrm{id}$. This means that γ and δ are mutually inverse homeomorphisms and so the preimages H and $\dot{H}$ are isomorphic.

The uniqueness of the smallest preimage and assertion 3) are checked in a similar manner. The theorem is proved.

This Theorem will be used for the proof of the following Theorems 2 and 3.

Now with the help of this Theorem we shall give a functional characterization of the Borelian cover.

Let $\{H, \eta\colon H \to T,\ T_S \mapsto H_S,\ \mathscr{C}_0(H)\}$ be a perfect saturated preimage of T lifting separable covering and having a completely normal base.

Let P and Q be subsets of H. If $P \setminus Q \not\supset H_S$ for any S we shall say that P is *almost contained in* Q and write $P \underset{\sim}{\subset} Q$.

Let $\{C_k\} \subset \mathscr{C}_0(H)$ be a finite covering of the space H and $\{B_k\} \subset \mathscr{B}(T)$ be a finite covering of the space T. The family $\{C_k, \eta^{-1}B_k\}$ will be called a *cohesive covering of the space* H if $\eta^{-1}B_k \underset{\sim}{\subset} C_k$ for any k. Note that for any cohesive covering $\{C_k, \eta^{-1}B_k\}$ the set $\cup(\eta^{-1}B_k \cap C_k)$ is dense in H.

LEMMA 10. *Let* $\{C_j, \eta^{-1}B_j\}$ *and* $\{C_k, \eta^{-1}B_k\}$ *be cohesive coverings. Then the family* $\{C_j \cap C_k, \eta^{-1}(B_j \cap B_k)\}$ *is a cohesive covering, too.*

PROOF. Assume that there exists a set S such that $H_S \subset \eta^{-1}(B_j \cap B_k) \setminus (C_j \cap C_k)$. As $H_S \cap C_j \neq \emptyset$ there exists a set $R \subset S$ such that $H_R \subset H_S \cap C_j$. This implies $H_R \subset \eta^{-1}B_k \setminus C_k$ but this is impossible. It follows from this contradiction that $\eta^{-1}(B_j \cap B_k) \underset{\sim}{\subset} C_j \cap C_k$.

Further for a natural number n the number $1/n$ will be denoted by u_n.

Let f and g be functions on H. The functions f and g will be called *equivalent* if for any n there exists a cohesive covering $\{C_k, \eta^{-1}B_k\}$ of H such that $|f(s)-g(s)| < u_n$ for any $s \notin \bigcup(\eta^{-1}B_k \setminus C_k)$. In this case we shall write $f \sim g$. It follows from the previous lemma that this relation "$\sim$" is indeed an equivalence relation.

Consider on H the set $C_0^*(H)$ of all functions $f \in C^*(H)$ such that

$$f^{-1}(]a, b[) \in \mathscr{C}_0(H)$$

for any open interval $]a, b[$. It follows from the theorem of Alexandrov ([11]) that $C_0^*(H)$ is a uniformly complete vector lattice and $\mathscr{C}_0(H) = \{\operatorname{coz} f \mid f \in C_0^*(H)\}$.

The set of all bounded Borel functions on T will be denoted by $B^*(T)$.

We have the following functional description of Borel determinedness.

LEMMA 11. *The following assertions are equivalent:*

a) *H is Borel determined;*

b) *for any function $f \in C_0^*(H)$ there is a (unique) Borel function $x \in B^*(T)$ such that $f \sim x \circ \eta$.*

PROOF. Denote $B^*(T)$ by X and $C_0^*(H)$ by Φ. Let H be Borel determined. Consider a function $0 \leqq f \in \Phi$. Divide an interval, containing the range of the function f, by points a_{mj} so that $a_{m,j+1} - a_{m,j} = u_m/6$. Then for any $C_{mj} \equiv f^{-1}(]a_{m,j-1}, a_{m,j+1}[)$ there exists a Borel set B'_{mj} such that $\eta^{-1}B'_{mj} \triangle C_{mj} \not\supset H_S$ for any S. Assume that there exists a point $t \notin \bigcup B'_{mj}$. As $H_t \cap C_{mj} \neq \emptyset$ for some j we have $H_t \subset C_{mj}$. So $H_t \subset C_{mj} \setminus \eta^{-1}B'_{mj}$ but this is impossible. This means that $\{B'_{mj}\}$ is a covering of T for any m. Consider the sets $B_{mj} \equiv B'_{mj} \setminus \bigcup\{B'_{mi} \mid i<j\}$. Then $\{C_{mj}, \eta^{-1}B_{mj}\}$ is a cohesive covering. Consider the Borel step function $x_m \equiv \sum a_{mj}\chi(B_{mj})$. Let $x_n \equiv \equiv \sum' a_{nk}\chi(B_{nk})$. Take a point $t \in T$. Then $t \in B_{mj} \cap B_{nk}$ for some j and k. So $H_t \subset \subset \eta^{-1}B_{mj}$ implies $H_t \subset C_{mj}$. Similarly $H_t \subset C_{nk}$. Take a point $s \in H_t$. Then we have $|x_m(t)-x_n(t)| = |a_{mj}-a_{nk}| \leqq |a_{mj}-f(s)| + |f(s)-a_{nk}| < u_m/3$ for $n \geqq m$. That is why there exists a Borel function x such that $|x(t)-x_n(t)| < 2u_n/3$. Let $s \notin P_n \equiv \equiv \bigcup((\eta^{-1}B_{nk}) \setminus C_{nk})$. Then $|f(s) - x \circ \eta(s)| < u_n$. Thus $f \sim x \circ \eta$.

Assume that there exists another function $x' \in X$ having this property. Then for any n there exists a cohesive covering $\{C_l, \eta^{-1}B_l\}$ such that $|f(s) - x' \circ \eta(s)| < u_n$ for any $s \notin Q_n \equiv \bigcup(\eta^{-1}B_l \setminus C_l)$. Take a point $t \in T$. Then $t \in B_{nk} \cap B_l$ for some indexes. Therefore $H_t \subset \eta^{-1}(B_{nk} \cap B_l)$. By virtue of Lemma 10 there exists a point $s \in H_t \cap \eta^{-1}(B_{nk} \cap B_l) \cap (C_{nk} \cap C_l)$. So $\eta s = t$. Besides $s \notin P_n \cup Q_n$. Hence we get $|x(t) - x'(t)| < 2u_n$. Consequently $x = x'$.

Now let C be a cozero-set from the base. Then $C = \operatorname{coz} f$ for some function $0 \leqq f \in \Phi$. Consider for the function f the corresponding Borel function $x \geqq 0$ such that $f \sim x \circ \eta$. Consider the Borel set $B \equiv \operatorname{coz} x$. Assume that $H_S \subset \eta^{-1}B \triangle C \equiv Q$. If $H_S \cap (C \setminus \eta^{-1}B) \neq \emptyset$ there exists a set $R \subset S$ such that $H_R \subset C \setminus \eta^{-1}B$. Consider the sets $C_n \equiv \{s \in H \mid f(s) > u_n\}$. For some n there exists a set $R_1 \subset R$ such that $H_{R_1} \subset C_n \cap H_R \subset C_n \setminus \eta^{-1}B$. But for this n there exists a cohesive covering $\{C_k, \eta^{-1}B_k\}$ such that $|f(s) - x \circ \eta(s)| < u_n$ for any $s \notin P_n \equiv \cup(\eta^{-1}B_k \setminus C_k)$. It is clear that $H_{R_1} \subset P_n$. As $R_1 \cap B_k \neq \emptyset$ for some k, there exists a compact set $R_2 \subset \subset R_1 \cap B_k$. So $H_{R_2} \subset \eta^{-1}B_k \setminus C_k$ but this is impossible. Consequently $H_S \subset \eta^{-1}B \setminus C$. Consider the sets $B_n \equiv \{t \in T \mid x(t) > u_n\}$. Then $S \cap B_n \neq \emptyset$ for some n. Hence there exists a compact set $R \subset S \cap B_n$. Therefore $H_R \subset \eta^{-1}B_n \setminus C$. This implies $H_R \subset P_n$. But as it was shown above this is impossible. From this contradiction we get that $\eta^{-1}B \triangle C \not\supset H_S$ for any S. So H is Borel determined. The lemma is proved.

PROPOSITION 1. *Let H be the Borelian cover of T. Then*

(a) *$\{H, \eta\colon H \to T, T_S \mapsto H_S, \mathscr{C}_0(H)\}$ is a perfect saturated preimage of T lifting separable covering and having a completely normal base;*

(b) *there is a bijection $r\colon x \mapsto f$ between the family $B^*(T)$ and the family $C_0^*(H)$ such that $f \sim x \circ \eta$;*

(c) *H as a preimage of T lifting separable covering is completely determined (up to isomorphism) by the properties* (a)—(b).

PROOF. Denote $B^*(T)$ by X and $C_0^*(H)$ by Φ. Let $0 \leqq x \in X$. Then there exist step functions $x_n \in X$ such that $|x(t) - x_n(t)| < u_n$ for any t and $x_n \equiv \sum a_k \chi(B_k)$ for some Borel partitions $\{B_k\}$ of T. Denote the set iB_k by U_k. Consider the functions $f_n \equiv \sum a_k \chi(U_k) \in \Phi$. Let $0 \leqq f \in \Phi$ be the uniform limit of the sequence f_m. By virtue of Corollary 2 of Lemma 3 the family $\{U_k, \eta^{-1}B_k\}$ is a cohesive covering. Let $s \notin P_n \equiv \cup(\eta^{-1}B_k \setminus U_k)$. Then $|f(s) - x \circ \eta(s)| < 3u_n$. Hence $f \sim x \circ \eta$.

Assume that there is another function $f' \in \Phi$ satisfying this condition, i.e. for any n there exists a cohesive covering $\{C_l, \eta^{-1}B_l\}$ such that $|f'(s) - x \circ \eta(s)| < u_n$ for any $s \notin Q_n \equiv \cup(\eta^{-1}B_l \setminus C_l)$. We can suppose that $\{B_l\}$ is a partition. Take a point $s \in \cup((\eta^{-1}(B_k \cap B_l)) \cap (U_k \cap C_l)) \equiv R_n$. Then $s \notin P_n \cup Q_n$ implies that $|f(s) - f'(s)| < < 4u_n$. By virtue of Lemma 10 and the density of the set R_n we conclude that this inequality is valid for any $s \in H$. As a result we get $f = f'$.

Thus the mapping $r\colon x \mapsto f$ is defined correctly. By virtue of Lemma 11 we get that this mapping is bijective.

Now let $\{\dot H, \dot\eta\colon \dot H \to T, T_S \mapsto \dot H_S, \mathscr{C}_0(\dot H)\}$ be a preimage of T with the properties from (a) and (b). Denote the vector lattice $C_0^*(\dot H)$ by $\dot\Phi$. It is easy to check that the mapping $\dot r\colon X \to \dot\Phi$ is an isomorphism of the vector lattices. By virtue of Lemma 11 $\dot H$ is Borel determined.

Let C be the cozero-set of a function $f = \dot r x$. Consider the functions $y \equiv \sup\{nx \wedge 1 \mid n\} \in X$ and $g \equiv \dot r y$. Then $g = \sup\{nf \wedge 1\}$ in $\dot\Phi$. Hence $g = \chi(\mathrm{cl}\, C)$. Consequently $U \equiv \mathrm{cl}\, C \in \Delta_0(\dot H)$. Thus $\dot H$ is σ-extremally disconnected. Let $C \cap \dot H_S = \emptyset$. Denote the $\mathrm{coz}\, x$ by B. It follows from the proof of Lemma 11 that $C \triangle \dot\eta^{-1}B \not\supset \dot H_R$ for any R. Assume that $B \cap S \neq \emptyset$ and take a compact subset R from this intersection. Then $\dot H_R \subset \dot\eta^{-1}B \cap \dot H_S \subset \dot\eta^{-1}B \setminus C$ but this is impossible. So $B \cap S = \emptyset$. As $B = \mathrm{coz}\, y$, $U = \mathrm{coz}\, g$ and $g = \dot r y$ then as above $U \triangle \dot\eta^{-1}B \not\supset \dot H_R$ for any R. Assume $U \cap \dot H_S \neq \emptyset$. By virtue of the saturatedness there exists a set $R \subset S$ such that $\dot H_R \subset \subset U \cap \dot H_S$. Then $\dot H_R \cap \dot\eta^{-1}B \neq \emptyset$. But this contradicts to $B \cap R = \emptyset$. This means that the preimage $\dot H$ is σ-disjoined.

Let G be an open set from T and $x \equiv \chi(G)$. Then $x = \sup\{f_r \in C^*(T) \mid f_r \leqq x\}$ in X. Consider the function $f = \dot r x$. Then $f = \sup\{\dot r f_r\} = \sup\{f_r \circ \dot\eta\} = \chi(U)$, where

$U \equiv \mathrm{cl}\, \dot{\eta}^{-1}G$. Hence $U \in \Delta_0(\dot{H})$. This means that $\dot{H}$ is lower extremally disconnected. By similar arguments as above it is verified that $\dot{H}$ is lower disjoined. On the strength of Theorem 1 we conclude that the preimages H and $\dot{H}$ are isomorphic. The proposition is proved.

This proposition also will be used in the sequel.

Maks some remarks. Substituting open set by cozero-sets in all the above definitions we can define the field $\mathscr{B}_\sigma(T)$ of all *Baire subsets of* T, the *Baire cover* H_σ *of* T, the notion of the *Baire determinedness*, the family $B_\sigma^*(T)$ of all bounded *Baire functions* and so on. In such manner we shall obtain the very similar theory of the Baire cover and its characterizations of the same form. In this case the notions of the lower extremally disconnectedness and the lower disjointedness are unnecessary.

In fact a more general form of these definitions is obtained by considering from the beginning a completely normal Alexandrov space [11] rather than the families of closed sets and of zero-sets in a completely regular space.

§ 2. Vector lattice of Borel functions

Let T be a completely regular space and $B^*(T)$ be the vector lattice of all bounded Borel functions on T. Let $u\colon C^*(T) \to B^*(T)$ be the canonical imbedding. For a countable set S consider the ideal $B_S^*(T) \equiv \{x \in B^*(T) | x(S)=0\}$. Then

$$\{B^*(T), u\colon C^*(T) \to B^*(T), C_S^*(T) \mapsto B_S^*(T)\}$$

is an extension of $C^*(T)$ inheriting separable decomposition. This extension will be called the *Borelian extension* of $C^*(T)$.

Let H be the Borelian cover of T and $\eta\colon H \to T$ be the canonical mapping. Let $\Phi \equiv C_0^*(H)$ be the vector lattice of functions on H defined in § 1. Consider the injective vector-lattice homomorphism $\varphi\colon C^*(T) \to \Phi$ such that $\varphi f \equiv f \circ \eta$. For a countable set S consider the ideal $\Phi_S \equiv \{f \in \Phi | f(H_S)=0\}$. Then

$$\{\Phi, \varphi\colon C^*(T) \to \Phi, C_S^*(T) \mapsto \Phi_S\}$$

is an extension of $C^*(T)$ inheriting separable decomposition.

Now let $\{X, u\colon C^*(T) \to X, C_S^*(T) \mapsto X_S\}$ be a vector-lattice extension of $C^*(T)$ inheriting separable decomposition. Identify $C^*(T)$ with its image in X.

Let $x \in X$ and $\{x_\xi\} \subset X$. The element x will be called the *d-supremum of the set* $\{x_\xi\}$ if $x \geqq x_\xi$ and for any X_S we have $\bar{x} = \sup \bar{x}_\xi$ in X/X_S. In this case we shall write $x = d\text{-}\sup x_\xi$. In a similar way the *d-infimum of the set* $\{x_\xi\}$ is defined.

Consider the sets $S_l(C^*(T), X) \equiv \{x \in X | \exists f_\xi \in C^*(T)\ (x = d\text{-}\sup f_\xi)\}$,

$$B_0(C^*(T), X) \equiv \{x - y | x, y \in S_l(C^*(T), X)\}$$

and $B(C^*(T), X) \equiv \bigcup \{B_\alpha(C^*(T), X) | \alpha < \omega_1\}$, where

$$B_\alpha(C^*(T), X) \equiv$$

$$\equiv \{x \in X | \exists \alpha_n < \alpha\, \exists \uparrow u_n \in B_{\alpha_n}(C^*(T), X)\, \exists \downarrow v_n \in B_{\alpha_n}(C^*(T), X) (x = d\text{-}\sup u_n = d\text{-}\inf v_n)\}.$$

The extension X will be called *Borel generated* if $X=B(C^*(T), X)$.

For the extensions $B^*(T)$ and Φ consider the mapping $r\colon B^*(T)\to\Phi$ from Proposition 1 of the previous paragraph.

PROPOSITION 2. *With respect to the mapping r the extensions $B^*(T)$ and Φ are isomorphic saturated Borel generated lower Dedekind complete σ-Dedekind complete lower component σ-component extensions of $C^*(T)$ inheriting separable decomposition.*

PROOF. Denote $B^*(T)$ by X and $B_S^*(T)$ by X_S. It can be verified that $r\circ u=\varphi$ and r is an isomorphism of the vector lattices. Let $x\in X$ and $f\equiv rx$. Consider the sets $B\equiv\operatorname{coz} x$ and $C\equiv\operatorname{coz} f$. As established in the proof of Lemma 11, $C\triangle\eta^{-1}B\not\supset H_R$ for any R. Let $x\in X_S$ and assume that $f\notin\Phi_S$. By virtue of the saturatedness there exists a set $R\subset S$ such that $H_R\subset C\cap H_S$. We can suppose that R is compact. Hence $H_R\cap\eta^{-1}B\neq\emptyset$ implies $R\cap B\neq\emptyset$ but this is impossible. Thus our assumption is false. Conversely, let $f\in\Phi_S$ and assume that $B\cap S\neq\emptyset$. Take a compact set $R\subset B\cap S$. Then $H_R\subset\eta^{-1}B\setminus C$ but this is impossible. Consequently, $x\in X_S$. Thus the extensions X and Φ are isomorphic.

Let Y be a proper component of X such that $Y^d\not\subset X_S$. Consider the nonempty set $P\equiv\{t\in T\mid\forall y\in Y\ (y(t)=0)\}$. Then $Y=\{x\in X\mid x(P)=0\}$. Consequently, $R\equiv P\cap S\neq\emptyset$. Therefore $X_S\cup Y\subset X_R$. This means that X is saturated.

Now verify that X is Borel generated. Let $S_l^*(T)$ denote the set of all bounded lower semicontinuous functions on T. Then $B^*(T)=\cup\{B_\alpha^*(T)\mid\alpha<\omega_1\}$ where $B_0^*(T)\equiv\{x-y\mid x, y\in S_l^*(T)\}$ and

$$B_\alpha^*(T)\equiv$$

$$\equiv\{x\in B^*(T)\mid\exists\alpha_n<\alpha\,\exists\uparrow u_n\in B_{\alpha_n}^*(T)\,\exists\downarrow v_n\in B_{\alpha_n}^*(T)\,\forall t\in T(x(t)=\sup u_n(t)=\inf v_n(t))\}.$$

Let $x\in S_l^*(T)$. Then $x(t)=\sup\{f_\xi(t)\}$ for some family $f_\xi\in C^*(T)$. It can be checked that $\bar{x}=\sup\bar{f}_\xi$ in any X/X_S. So $x=d\text{-}\sup f_\xi$. Thus $B_0^*(T)\subset B_0(C^*(T), X)$. In a similar way it is checked that $B_\alpha^*(T)\subset B_\alpha(C^*(T), X)$. As the rest of the properties of X are well-known the proposition is proved.

Further uniqueness is understood up to isomorphism.

THEOREM 2. (1) *$B^*(T)$ is the unique largest of all the saturated Borel generated extensions of $C^*(T)$ inheriting separable decomposition;*

(2) *$B^*(T)$ is the unique smallest of all the filled lower Dedekind complete σ-Dedekind complete lower component σ-component extensions of $C^*(T)$ inheriting separable decomposition and moreover $B^*(T)$ is the unique universal among all such extensions;*

(3) *$B^*(T)$ is the unique saturated Borel generated lower Dedekind complete σ-Dedekind complete lower component σ-component extension of $C^*(T)$ inheriting separable decomposition.*

PROOF. Let $\{X, u\colon C^*(T)\to X, C_S^*(T)\mapsto X_S\}$ be an extension having the properties from 1). On the strength of Yosida's theorem ([12]) there is unique compact H_0 such that the vector lattice X is isomorphic to the vector lattice $C(H_0)$ relative to an isomorphism r_0. Then the mapping u generates a unique surjective continuous mapping η_0: $H_0 \to \beta T$ such that $r_0uf=f'\circ\eta_0$, where f' denotes the extension of a function $f\in C^*(T)$ on βT.

Consider the space $H\equiv\eta_0^{-1}T$ and the perfect mapping η: $H\to T$ which is the restriction of η_0. Consider the vector lattice Φ, consisting of the restrictions on H of all functions from $C(H_0)$, the homomorphism $r:X\to\varphi$ such that $rx\equiv r_0x|H$, and the homomorphism φ: $C^*(T)\to\Phi$ such that $\varphi f\equiv f\circ\eta$.

For a countable set S consider the ideals $\Phi_{0_S}\equiv r_0X_S$ and $\Phi_S\equiv rX_S$ and the closed subsets $H_{0_S}\equiv\{s\in H_0|\forall f\in\Phi_{0_S}(f(s)=0)\}\neq\emptyset$ and $H_S\equiv H_{0_S}\cap H$. Then $\cup H_{0_S}$ is dense in H_0 and $\eta_0H_{0_S}=\mathrm{cl}\,T_S$. It is clear that $S_1\subset S_2$ implies $H_{S_1}\subset H_{S_2}$. Let R be a compact subset of S. Then $H_R=H_{0_R}$. It follows from this fact that $H_S\neq\emptyset$ for any S.

Let $f\geqq 0$ be a function from $C(H_0)$ such that $f(H_{0_S})=0$. Consider the functions $f_k\equiv(f-u_k\mathbf{1})\vee 0$. From the property $H_{0_S}\cap\mathrm{cl}\,\mathrm{coz}\,f_k=\emptyset$ we conclude that $f_k\in\Phi_{0_S}$. This implies that f belongs to this set, too. Thus $\Phi_{0_S}=\{f\in C(H_0)|\,f(H_{0_S})=0\}$.

Let C be the cozero-set of a function $f\in C(H_0)$ such that $C\cap H_{0_S}\neq\emptyset$. Represent S in the form $S=\cup R_k$ for some compact subsets R_k. As X is filled we have $f\in\Phi_{0_{R_k}}$ for some k. Therefore $C\cap H_S\supset C\cap H_{0_{R_k}}\neq\emptyset$. This means that H_S is dense in H_{0_S}. As a consequence we get $\Phi_S=\{f\in\Phi|\,f(H_S)=0\}$ and $\eta H_S=T_S$.

Besides we established that H is dense in H_0. Hence the triplet

$$\{\Phi,\varphi\colon C^*(T)\to\Phi,\ C_S^*(T)\mapsto\Phi_S\}$$

is an extension isomorphic to the initial one.

In addition we get that $\cup H_S$ is dense in H. Consequently, H is the preimage of T lifting separable covering and having the completely normal base $\mathscr{C}_0(H)\equiv\equiv\{\mathrm{coz}\,f|\,f\in\Phi\}$.

Let G be an open set in H and $G\cap H_S\neq\emptyset$. Take a proper regular closed set $F\subset G$ such that $H_S\cap\mathrm{int}\,F\neq\emptyset$. Consider the proper component $Y\equiv\{f\in\Phi|\,f(F)=0\}$. As $Y^d\not\subset\Phi_S$ we get by virtue of the saturatedness that there exists an ideal Φ_R containing the set $\Phi_S\cup Y$. This means that $H_R\subset H_S\cap G$. Thus H is a saturated preimage.

Now let f be an element from $S_l(C^*(T),\Phi)$ and $C\equiv\mathrm{coz}\,f$. Then $f=d\text{-}\sup\{f_\xi\circ\eta\}$. Consider the lower semi-continuous function x on T, such that $x(t)\equiv\sup\{f_\xi(t)\}$, and the Borel set $B\equiv\mathrm{coz}\,x$. Assume that $H_S\subset C\triangle\eta^{-1}B$ for some S. If $H_S\cap\cap(C\setminus\eta^{-1}B)\neq\emptyset$ then by virtue of the saturatedness there exists a set $R\subset S$ such that $H_R\subset C\setminus\eta^{-1}B$. This means that $f_\xi\in C_R^*(T)$ and $f\notin\Phi_R$. But this contradicts to the equality $\bar f=\sup\{\overline{\varphi f_\xi}\}$ in Φ/Φ_R. Hence $H_S\subset\eta^{-1}B\setminus C$. This implies that $f_\xi\notin C_S^*(T)$ for some ξ and $f\in\Phi_S$, but this is also impossible. Thus such a set S does not exist.

Further we shall proceed by induction. Assume that for any $\beta<\alpha$ and for any $f\in B_\beta(C^*(T),\Phi)$ there exists a Borel set B such that $\mathrm{coz}\,f\triangle\eta^{-1}B\not\supset H_S$ for any S. Let f be a function from $B_\alpha(C^*(T),\Phi)$ and $C\equiv\mathrm{coz}\,f$. Then $f=d\text{-}\sup u_k=d\text{-}\inf v_k$ for some sequences $u_k,v_k\in B_{\alpha_k}(C^*(T),\Phi)$. Consider for the sets $C_k\equiv\mathrm{coz}\,u_k$ the corresponding Borel sets B_k and the set $B\equiv\cup B_k$. Assume that $H_S\subset C\triangle\eta^{-1}B$ for some S. If $H_S\cap(C\setminus\eta^{-1}B)\neq\emptyset$ then there exists a set $R\subset S$ such that $H_R\subset\subset C\setminus\eta^{-1}B$. This means that $f\notin\Phi_R$. Therefore $u_k\notin\Phi_R$ for some k. Hence $C_k\cap\cap H_R\neq\emptyset$ implies that there exists a set $R_1\subset R$ such that $H_{R_1}\subset C_k\setminus\eta^{-1}B_k$. But this is impossible. Thus $H_S\subset\eta^{-1}B\setminus C$. So there exists a compact set $R\subset B_k\cap S$ for some k. Then $H_R \subset \eta^{-1} B_k \setminus C_k$ but this is also impossible. From this contradiction we conclude that such a set S does not exist.

As a result we get that for any cozero-set C from the chosen base there exists a corresponding Borel set B. This means that the preimage H is Borel determined.

Now let $\{\dot{X}, \dot{u}: C^*(T) \to \dot{X}, C_S^*(T) \mapsto \dot{X}_S\}$ be an extension having the properties from 2). Consider, as it was done above, the isomorphic extension

$$\{\dot{\Phi}, \dot{\varphi}: C^*(T) \to \dot{\Phi}, C_S^*(T) \mapsto \dot{\Phi}_S\}$$

for the corresponding preimage $\{\dot{H}, \dot{\eta}: \dot{H} \to T, T_S \mapsto \dot{H}_S, \mathscr{C}_0(\dot{H})\}$.

Let $S = \bigcup S_k$ for some sequence of subsets S_k. Then $\dot{\Phi}_S = \bigcap \dot{\Phi}_{S_k}$ implies that $\bigcup \dot{H}_{S_k}$ is dense in $\dot{H}_S$. This means that the preimage $\dot{H}$ is filled. Let G be an open set from T. Consider the family $\{f_\xi\}$ consisting of all continuous functions which are smaller than the characteristic function of the set G. Consider the function $f \equiv \sup \{\dot{\varphi} f_\xi\}$. Then $f(s) = 1$ for any $s \in \dot{\eta}^{-1} G$ and $f(s) = 0$ for any $s \notin U \equiv \mathrm{cl}\, \dot{\eta}^{-1} G$. From the continuity of the function f we conclude that $f = \chi(U)$. So $U \in \Delta_0(\dot{H})$. Thus the preimage $\dot{H}$ is lower extremally disconnected.

Let $\dot{\eta}^{-1} G \cap \dot{H}_S = \emptyset$. Then $f_\xi \in C_S^*(T)$ implies that $f \in \dot{\Phi}_S$. Therefore $U \cap \dot{H}_S = \emptyset$. This means that $\dot{H}$ is lower disjoined.

Let C be the cozero-set of a function $f \in \dot{\Phi}$. Consider the function $g \equiv \sup \{nf \wedge \mathbf{1}|n\}$. Then $g = \chi(U)$ for the set $U \equiv \mathrm{cl}\, C$. So $U \in \Delta_0(\dot{H})$. Therefore $\dot{H}$ is σ-extremally disconnected. If $C \cap \dot{H}_S = \emptyset$ then $f \in \dot{\Phi}_S$ implies $g \in \dot{\Phi}_S$. Hence $U \cap \dot{H}_S = \emptyset$. Thus $\dot{H}$ is σ-disjoined.

On the strength of Theorem 1 there exists a mapping $\gamma: \dot{H} \to H$ such that $\dot{H}$ is larger than H relative to γ.

Let C be a cozero-set from the base on H. Check that $\mathrm{cl}\, \gamma^{-1} C = \alpha C$ where $\alpha: \mathscr{C}_0(H) \to \Delta_0(\dot{H})$ is the lattice homomorphism from the proof of Theorem 1. Let $s \in \gamma^{-1} C$. Take a set C_1 from the base such that $\gamma s \notin \mathrm{cl}\, C_1$ and $C \cup C_1 = H$. Then $\dot{H} = \alpha C \cup \alpha C_1$. If we assume that $s \in \alpha C_1$ then we get $\gamma s \in \mathrm{cl}\, C_1$. As this is false we have $s \in \alpha C$. Thus $\gamma^{-1} C \subset \alpha C$. Further consider cozero-sets C_n from the base such that $C = \bigcup C_m$ and $\mathrm{cl}\, C_m \subset C$. Consider the Borel sets $B_m \equiv k C_m$ and the set $B \equiv \bigcup B_S$. Assume that $C \triangle \eta^{-1} B \supset H_S$ for some S. If $H_S \cap (C \setminus \eta^{-1} B) \neq \emptyset$ then there exists a set $R \subset S$ such that $H_R \subset C_m \setminus \eta^{-1} B$ for some m but this is impossible. Therefore $H_S \subset \eta^{-1} B \setminus C$. This implies that there exists a compact set $R \subset S \cap B_m$ for some m. Then $H_R \subset \eta^{-1} B_m \setminus C$ but this is also impossible. From this contradiction we conclude that such a set S does not exist. Hence $B = kC$. It was established in the proof of Theorem 1 that $\mathrm{i}B = \mathrm{cl} \bigcup \mathrm{i} B_m$. Consequently, $\alpha C = \mathrm{cl} \bigcup \alpha C_m$. If $s \in \alpha C_m$ then $\gamma s \in \mathrm{cl}\, C_m \subset C$. So $\alpha C = \mathrm{cl}\, \gamma^{-1} C$.

Now let $f \geqq 0$ be a function from Φ. Divide an interval, containing the range of f, by points a_j so that $a_{j+1} - a_j = u_m/2$. Consider the sets $C_j \equiv f^{-1}(]a_{j-1}, a_{j+1}[)$ and $U_j \equiv \mathrm{cl}\, \gamma^{-1} C_j = \alpha C_j$. It was established above that there exist functions $g_j \in \dot{\Phi}$ which are the characteristic functions of the sets αC_j. Consider the step function $g_m \equiv \sup \{a_{j-1} g_j\}$. It is clear that $0 \leqq f \circ \gamma(s) - g_m(s) \leqq u_m$ for any s. Therefore $f \circ \gamma \in \dot{\Phi}$.

This means that we can define correctly the injective vector-lattice homomorphism $v: \Phi \to \dot{\Phi}$ by setting $vf \equiv f \circ \gamma$. Then $\dot{\varphi} = v \circ \varphi$. Let $f \in \Phi_S$. Then $(vf)(\dot{H}_S) = 0$

implies $vf \in \dot{\Phi}_S$. Thus the extension $\dot{\Phi}$ is larger than the extension Φ. This fact is valid for the initial extensions $\dot{X}$ and X, too.

Now let Φ be the extension from Proposition 2 isomorphic to the Borelian extension $B^*(T)$. As Φ has the properties from (1) and (2) simultaneously we get as a result that Φ is the largest of all the extensions with the properties from (1) and the smallest of all the extensions with the properties from (2).

Let $\dot{X}$ be some other largest extension of $C^*(T)$. Consider, as it was done above, the isomorphic extension $\{\dot{\Phi}, \dot{\varphi}\colon C^*(T) \to \dot{\Phi}, C_S^* \mapsto \dot{\Phi}_S\}$ for the preimage $\{\dot{H}, \dot{\eta}\colon \dot{H} \to T, T_S \mapsto \dot{H}_S, \mathscr{C}_0(\dot{H})\}$. Take some mapping $w\colon \Phi \to \dot{\Phi}$ such that $\dot{\Phi}$ is larger than Φ relative to w. Define the surjective perfect mapping $\delta\colon \dot{H} \to H$ by setting $\delta s \equiv \bigcap \{\mathrm{cl}\,\mathrm{coz} f \cap \eta^{-1}\dot{\eta}s \mid s \in \mathrm{coz}\, wf\}$. Then $\eta \circ \delta = \dot{\eta}$. Check that $wf = f \circ \delta$ for any function $0 \leqq f \in \Phi$. Assume that there exists a point s such that $(wf)(s) \neq (f \circ \delta)(s)$. If $(wf)(s) > (f \circ \delta)(s)$ then we shall consider the function $g \equiv f$ otherwise $g \equiv -f$. Denote the number $((wg)(s) + (g \circ \delta)(s))/2$ by a. Consider the function $h \equiv (g - a1) \vee 0$. Take a neighbourhood G of s such that $(wg)(t) > a$ for any $t \in G$. Also take a neighbourhood U of the point δs such that $g(r) < a$ for any $r \in U$. Then $U \subset H \setminus \mathrm{coz}\, h$ and $G \cap \delta^{-1} U \subset \mathrm{coz}\, wh$. Therefore $\delta s \notin \mathrm{cl}\,\mathrm{coz}\, h$ and $\delta s \in \mathrm{cl}\,\mathrm{coz}\, h$ but this is impossible. From this contradiction we conclude that such a point s does not exist.

Check that $\delta \dot{H}_S \subset H_S$. Assume that there exists a point $s \in \delta \dot{H}_S \setminus H_S$. Take a function $f \in \Phi_S$ such that $s \in \mathrm{coz} f$. Then for some point $t \in \dot{H}_S$ such that $s = \delta t$ we get $(wf)(t) \neq 0$. But on the other hand $wf \in \dot{\Phi}_S$ implies $(wf)(t) = 0$. It follows from this contradiction that this inclusion is valid.

Now consider the mapping $v\colon \dot{\Phi} \to \Phi$ defined above. Let s be a point from H_S. By virtue of the saturatedness of the Borelian cover we have $s = \bigcap \{H_R\}$. Then $\delta \gamma s \in \bigcap \{H_R\} = s$. From this fact we conclude that $\delta \gamma s = s$ for any point $s \in H$. Therefore $(wvf)(s) = f(s)$. Thus v and w are mutually inverse isomorphisms of vector lattices. So the extensions Φ and $\dot{\Phi}$ are isomorphic.

The uniqueness of the smallest extension and assertion (3) are checked in a similar way. The theorem is proved.

Make some remarks. In a similar manner we obtain the characterization of the *Baire extension* $B_\sigma^*(T)$ consisting of all bounded Baire functions on T. In this case the notions of the lower Dedekind completeness and the lower componentness are unnecessary.

§ 3. Lattice ring of Borel functions

Let T be a completely regular space and $B^*(T)$ be the f-ring of all bounded Borel functions on T. Let $u\colon C^*(T) \to B^*(T)$ be the canonical imbedding. For a countable set S consider the f-ring ideal $B_S^*(T) \equiv \{x \in B^*(T) \mid x(S) = 0\}$. Then $\{B^*(T), u\colon C^*(T) \to B^*(T), C_S^*(T) \mapsto B_S^*(T)\}$ is an extension of $C^*(T)$ inheriting separable decomposition. This extension will be called the *Borelian extension* of $C^*(T)$.

Let H be the Borelian cover of T and $\eta\colon H \to T$ be the canonical mapping. Let $\Phi \equiv C_0^*(H)$ be the f-ring of functions on H defined in § 1. Consider the injective f-ring homomorphism $\varphi\colon C^*(T) \to \Phi$ such that $\varphi f \equiv f \circ \eta$. For a countable set S consider the f-ring ideal $\Phi_S \equiv \{f \in \Phi | f(H_S)=0\}$. Then $\{\Phi, \varphi: C^*(T) \to \Phi, C_S^*(T) \mapsto \Phi_S\}$ is an extension of $C^*(T)$ inheriting separable decomposition.

Now let $\{X, u: C^*(T) \to X, C_S^*(T) \mapsto X_S\}$ be an f-ring extension of $C^*(T)$ inheriting separable decomposition. Define the set $B(C^*(T), X)$ as it has been done in § 2. The extension X will be called *Borel generated* if $X=B(C^*(T), X)$.

For the extensions $B^*(T)$ and Φ consider the mapping $r: B^*(T) \to \Phi$ from Proposition 1 of § 1.

Proposition 3. *With respect to the mapping r the extensions $B^*(T)$ and Φ are isomorphic saturated Borel generated lower continuing σ-continuing lower segment σ-segment extensions of $C^*(T)$ inheriting separable decomposition.*

Proof. Denote $B^*(T)$ by X and $B_S^*(T)$ by X_S. It can be verified that $r \circ u = \varphi$ and r is an isomorphism of the f-rings. It has been checked in the proof of Proposition 2 that $x \in X_S$ iff $rx \in \Phi_S$. Therefore the extensions X and Φ are isomorphic.

In just the same way as in the proof of Proposition 2 it is checked that X is saturated and Borel generated.

Let Y be a ring ideal in the ring $C^*(T)$ and $g \in \mathrm{Hom}^*_{C^*(T)}(Y, C^*(T) \cap Y^{**})$. Let $y_1, y_2 \in Y$ and $t \in \mathrm{coz}\, y_1 \cap \mathrm{coz}\, y_2$. Then $(gy_1)(t)/y_1(t) = (gy_2)(t)/y_2(t)$. Consequently we can define correctly the Borel function $z \in X$ by setting $z(t) \equiv (gy)(t)/y(t)$ for any $y \in Y$ and any $t \in \mathrm{coz}\, y$ and $z(t) \equiv 0$ for any $t \notin G \equiv \bigcup \{\mathrm{coz}\, y | y \in Y\}$.

As $z \in Y^{**}$ we can define correctly the homomorphism $h \in \mathrm{Hom}^*_X(X, Y^{**})$ by setting $hx \equiv xz$. Let $y \in Y$ and $t \in G$. Then there exists a $y_1 \in Y$ such that $t \in \mathrm{coz}\, y_1$. Therefore $y_1(t)(hy)(t) = y(t)(gy_1)(t) = y_1(t)(gy)(t)$ implies $(hy)(t) = (gy)(t)$. As hy and gy belong to Y^{**} we have $(hy)(t) = 0 = (gy)(t)$ for any $t \notin G$. This means that $hy = gy$. Thus X is lower continuing.

Now let g and h be the homomorphisms from the definition of the lower segment and $g(Y) \subset X_S$. Let $x \in X$ and $t \in S \cap G$. Then $t \in \mathrm{coz}\, y$ for some $y \in Y$ implies $y(t)(hx)(t) = x(t)(gy)(t) = 0$ and hence $(hx)(t) = 0$. If $t \in S \setminus G$ then $(hx)(t) = 0$ because of $hx \in Y^{**}$. Consequently, $hx \in X_S$. This means that X_S is a lower segment of X.

Let Y be a ring ideal in the ring X, generated by a countable set $\{y_n\}$, and $g \in \mathrm{Hom}^*_X(Y, Y^{**})$. Consider the Borel sets $B_n \equiv \mathrm{coz}\, y_n$ and $B \equiv \bigcup B_n$. Define the Borel function $z \in X$ by setting $z(t) \equiv (gy_n)(t)/y_n(t)$ for any n and any $t \in B_n$ and $z(t) \equiv 0$ for any $t \notin B$. Let $x \in Y^*$. Then for a point $t \in B$ we have $x(t)z(t) = (g(xy_n))(t)/y_n(t) = 0$. Therefore $xz = 0$ implies $z \in Y^{**}$. That is why we can define correctly the homomorphism $h \in \mathrm{Hom}^*_X(X, Y^{**})$ by setting $hx \equiv xz$. Let $y \in Y$ and $t \in B$. Then $t \in B_n$ for some n. Hence $y_n(t)(hy)(t) = y(t)(gy_n)(t) = y_n(t)(gy)(t)$ implies $(hy)(t) = (gy)(t)$. Let x be the characteristic function of the set $T \setminus B$. Since an arbitrary element of the ideal Y is of the form $\sum x_{n_1 \dots n_k} y_{n_1} \dots y_{n_k}$ we conclude that $x \in Y^*$. That is why $(hy)(t) = 0 = (xgy)(t) = (gy)(t)$ for any $t \in T \setminus B$. Hence $hy = gy$. Thus X is σ-continuing.

As earlier it is checked that any X_S is a σ-segment of X. The proposition is proved.

Further uniqueness is understood up to isomorphism.

Theorem 3. 1) *$B^*(T)$ is the unique largest of all the saturated Borel generated extensions of $C^*(T)$ inheriting separable decomposition;*

2) *$B^*(T)$ is the unique smallest of all the filled lower continuing σ-continuing lower segment σ-segment extensions of $C^*(T)$ inheriting separable decomposition and moreover $B^*(T)$ is the unique universal among all such extensions;*

3) *$B^*(T)$ is the unique saturated Borel generated lower continuing σ-continuing lower segment σ-segment extension of $C^*(T)$ inheriting separable decomposition.*

PROOF. Let $\{X, u\colon C^*(T)\to X, C_S^*(T)\mapsto X_S\}$ be an extension having the properties from 1). On the strength of Johnson's theorem ([13]) there is a unique compact H_0 such that the f-ring X is isomorphic to the f-ring $C(H_0)$ relative to an isomorphism r_0. Further by completely the same arguments as in the proof of Theorem 2 we obtain the preimage $\{H, \eta\colon H\to T, T_S\to H_S, \mathscr{C}_0(H)\}$ of T and the corresponding extension $\{\Phi, \varphi\colon C^*(T)\to\Phi, C_S^*(T)\mapsto\Phi_S\}$ isomorphic to the initial one.

In just the same way as in the proof of Theorem 2 it is established that the preimage H is saturated and Borel determined.

Now let $\{\dot{X}, \dot{u}\colon C^*(T)\to\dot{X}, C_S^*(T)\mapsto\dot{X}_S\}$ be an extension having the properties from 2). Consider the isomorphic extension $\{\dot{\Phi}, \dot{\varphi}\colon C^*(T)\to\dot{\Phi}, C_S^*(T)\mapsto\dot{\Phi}_S\}$ for the corresponding preimage $\{\dot{H}, \dot{\eta}\colon \dot{H}\to T, T_S\mapsto\dot{H}_S, \mathscr{C}_0(\dot{H})\}$. Then the preimage $\dot{H}$ is filled.

Let G be an open set from T. Denote the set $\dot{\eta}^{-1}G$ by V. Consider the ring $R\equiv\dot{\varphi}C^*(T)$ and the ring ideal $Y\equiv\{y\in R|\operatorname{coz} y\subset V\}$ of the ring R. Define the homomorphism $g\in\operatorname{Hom}_R^*(Y, R\cap Y^{**})$ by setting $gy\equiv y$. Then there exists a bounded $\dot{\Phi}$-module homomorphism $h\colon \dot{\Phi}\to Y^{**}$ extending g. Consider the function $u\equiv h\mathbf{1}\in\dot{\Phi}$ and the set $U\equiv\operatorname{cl} V$. It is clear that $u(\dot{H}\setminus U)=0$. Let $s\in V$. Then $s\in\operatorname{coz} y$ for some $y\in Y$. Therefore $y(s)u(s)=(gy)(s)=y(s)$ implies $u(s)=1$. Since the function u is continuous we conclude that $u=\chi(U)$ and $U\in\Delta_0(\dot{H})$. So the preimage $\dot{H}$ is lower extremally disconnected.

Let $V\cap\dot{H}_S=\emptyset$. Then $g(Y)\subset\dot{\Phi}_S$ implies $u\in\dot{\Phi}_S$. Therefore $U\cap\dot{H}_S=\emptyset$. Thus the preimage $\dot{H}$ is lower disjoined.

Let C be the cozero-set of a function $v\in\dot{\Phi}$. Consider the principal ideal $Y\equiv\{fv|f\in\dot{\Phi}\}$ of the ring $\dot{\Phi}$. Define the bounded $\dot{\Phi}$-module homomorphism $g\colon Y\to Y^{**}$ by setting $gy\equiv y$. Then there exists a bounded $\dot{\Phi}$-module homomorphism $h\colon \dot{\Phi}\to Y^{**}$ extending g. Consider the function $u\equiv h\mathbf{1}\in\dot{\Phi}$ and the set $U\equiv\operatorname{cl} C$. It is clear that $u(\dot{H}\setminus U)=0$. Let $s\in C$. Then $v(s)u(s)=v(s)$ implies $u(s)=1$. This means that $u=\chi(U)$ and $U\in\Delta_0(\dot{H})$. So the preimage $\dot{H}$ is σ-extremally disconnected.

It is checked as above that $\dot{H}$ is σ-disjoined.

Further the proof goes on in exactly the same way as the proof of Theorem 2.

Make some remarks. In a similar manner we obtain the characterization of the Baire extension $B_\sigma^*(T)$ consisting of all bounded Baire functions on T. In this case the properties to be lower continuing and lower segment are unnecessary.

REFERENCES

[1] GORDON, H., The maximal ideal space of a ring of measurable functions, *Amer. J. Math.* **88** (1966), 827—843. *MR* 34 # 1838.

[2] SENTILLES, D., An L^1-space for Boolean algebras and semireflexivity of spaces $L^\infty(X, \Sigma, \mu)$, *Trans. Amer. Math. Soc.* **226** (1977), 1—37. *MR* 55 # 3774.

[3] Graves, W. H. and Sentilles, D., The extension and completion of the universal measure and the dual of the space of measures, *J. Math. Anal. Appl.* **68** (1979), 228—264. *MR* **80j**: 28014.
[4] Graves, W. H., On the theory of vector measures, *Mem. Amer. Math. Soc.* **12** (1977), no. 195. *MR* 57 # 7167c.
[5] Zaharov, V. K., Some perfect preimages connected with extensions of the family of continuous functions, *Topology, theory and applications* (Eger, 1983), Colloq. Math. Soc. J. Bolyai, 41, North-Holland, Amsterdam—New York, 1985. *MR* **87m**: 54043.
[6] Kuratowski, K., *Topology*, I—II, Państwowe Wydawnictwo Naukowe, Warsaw, 1966, 1968. *MR* 36 # 840; **41** # 4467.
[7] Arhangel'skiĭ, A. V. and Ponomarev, V. I., *Fundamentals of general topology through problems and exercises*, Nauka, Moscow, 1974 (in Russian). *MR* **56** # 3781; *Zbl* **287** # 54001.
[8] Sikorski, R., *Boolean algebras*, Berlin, 1964. *MR* **31** # 2178.
[9] Semadeni, Z., *Banach spaces of continuous functions*, Vol. 1, Monografie Matematyczne, Tom 55, PWN-Polish Scientific Publishers, Warsaw, 1971. *MR*. **45** # 5730.
[10] Fuchs, L., *Partially ordered algebraic systems*, Pergamon Press, Oxford, 1963. *MR* **30** # 2090.
[11] Alexandrov, A. D., Additive set-functions in abstract spaces I—III, *Rec. Math. [Mat. Sb.]* **8** (50) (1940), 307—348; **9** (51) (1941), 563—628; **13** (55) (1943), 169—238. *MR* **2**—315; **3**—207; **6**—275.
[12] Yosida, K., On vector lattice with a unit, *Proc. Imp. Acad. Tokyo* **17** (1941), 121—124. *MR* **3**—210.
[13] Johnson, D. G., On a representation theory for a class of Archimedean lattice-ordered rings, *Proc. London Math. Soc.* (3) **12** (1962), 207—225. *MR* **25** # 5082.

(Received May 23, 1985)

EÖTVÖS LORÁND TUDOMÁNYEGYETEM
TERMÉSZETTUDOMÁNYI KAR
ANALÍZIS TANSZÉK
MÚZEUM KRT. 6—8
H—1088 BUDAPEST
HUNGARY

Current address:

LENINGRAD INSTITUTE OF TEXTILE AND LIGHT INDUSTRY
HERZENA 18
SU—191065 LENINGRAD
SOVIET UNION

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 429—436

BANACH LIMITS AND RELATED MATRIX TRANSFORMATIONS

R. ÇOLAK and Ö. ÇAKAR

Summary

In this paper we establish a relationship between sublinear functionals which generate Banach limits and strongly regular matrices. Then, defining a sublinear functional on m, we show that this functional generates Banach limits and dominates all Banach limits. Furthermore, we give the relation between any sublinear functional defined on m and Banach limits by means of this functional. Finally we give the relations between sublinear functionals which generate Banach limits and dominate all Banach limits. It may be remarked that the elements of $\{m, \Omega\}$, for a particular sequence, are not necessarily unique, where Ω is a sublinear functional on m and $\{m, \Omega\}$ is the set of all linear functionals L defined on m such that $L(x) \leqq \Omega(x)$ for all $x \in m$. Therefore, we also examine the equality of the limits.

1. Introduction

Let N be the set of natural numbers, and let s, m, c and c_0 be the linear spaces of all real, bounded, convergent and null sequences $x=(x_k)$, respectively. We write

$$m_0 = \{x \in m\colon \sup_{n \in N} \Big|\sum_{i=1}^{n} x_i\Big| < \infty\}.$$

The linear spaces m, m_0, c and c_0 are normed by $\|x\| = \sup_{k \in N} |x_k|$.

We define $S\colon m \to R$ by $S(x) = \sup_k x_k$, $L^*\colon m \to R$ by $L^*(x) = \limsup_{k \to \infty} x_k$ and $\sigma\colon m \to m$ by $(\sigma x)_k = x_{k+1}$. Let m^* be the algebraic dual of m.

A Banach limit L is an element of m^* satisfying the conditions, [1],

(i) $x_k \geqq 0 \quad (k = 1, 2, \ldots) \Rightarrow L(x) \geqq 0$,

(ii) $L(e) = 1 \quad$ where $\quad e = (1, 1, \ldots)$,

(iii) $L(\sigma x) = L(x)$.

The condition (iii) is equivalent to say that L is σ-invariant and σ is called a shift operator. We denote the set of all Banach limits on m by $\mathcal{B}$ and the set of all linear functionals L defined on m such that $L(x) \leqq \Omega(x)$ for all $x \in m$ — where Ω is a sublinear functional on m — by $\{m, \Omega\}$; i.e.,

$$\{m, \Omega\} = \{L \in m^*\colon L(x) \leqq \Omega(x) \text{ for all } x \in m\}.$$

1980 *Mathematics Subject Classification* (1985 Revision). Primary 40CXX; Secondary 46AX.
Key words and phrases. Banach limits, matrix transformation to generate Banach limits, to dominate all Banach limits.

Let V be a sublinear functional on m. We say that (i) V generates Banach limits if $\{m, V\}\subset\mathscr{B}$ and (ii) V dominates all Banach limits if $\mathscr{B}\subset\{m, V\}$. Thus, if V both generates Banach limits and dominates all Banach limits, then $\{m, V\}=\mathscr{B}$ [9].

Let $A=(a_{nk})$ be an infinite matrix of real numbers. We write $A_n(x)=\sum_{k=1}^{\infty} a_{nk}x_k$ if the series converges for each $n\in N$, where $x=(x_k)\in s$. In this case we write $Ax=(A_n(x))$. A sequence (x_k) is called A-summable to a, if $A_n(x)=\sum_{k=1}^{\infty} a_{nk}x_k$ exists for all $n\in N$ and $A_n(x)\to a$ as $n\to\infty$ and it is written $A\text{-}\lim x_k=a$ or $\lim Ax=a$. We denote the convergence field of a matrix A by c_A, i.e., $c_A=\{x\colon Ax\in c\}$. We recall that the matrix A is called regular if $A\colon c\to c$ and $\lim A_n(x)=\lim x_k$. If the matrix A is regular then we say that A is almost positive if and only if $\lim_{n\to\infty}\sum_{k=1}^{\infty}|a_{nk}|=1$ [9]. A matrix A is called strongly regular if it is regular and $\lim_{n\to\infty}\sum_{k=1}^{\infty}|a_{nk}-a_{n,k+1}|=0$ ([6], [8]).

2. Banach limits and strongly regular matrices

LEMMA 1. *Let $V\colon m\to R$ be a sublinear functional. Then V generates Banach limits if and only if there exist strongly regular positive matrices A and B such that*

$$\lim_{n\to\infty} B_n(x) \leqq -V(-x) \leqq V(x) \leqq \lim_{n\to\infty} A_n(x) \tag{2.1}$$

for each $x\in m$.

PROOF. Since V is sublinear then by the Hahn—Banach extension theorem we can conclude that $\{m, V\}\neq\emptyset$.

Now suppose that V generates Banach limits. Then $V(x)\leqq W(x)$ for all $x\in m$ where $W(x)=\sup\{L(x)\colon L\in\mathscr{B}\}$, [9]. If we consider Lemma 1 in [5], p. 498 and the inequality $V(x)\leqq W(x)$ then it follows that there exist strongly regular positive matrices A and B such that

$$\lim_{n\to\infty} B_n(x)\leqq -V(-x)\leqq V(x)\leqq \lim_{n\to\infty} A_n(x)$$

holds. Conversely, if there exist strongly regular positive matrices A and B satisfying (2.1) and if $L\in\{m, V\}$ then the following inequalities are satisfied:

$$\lim_{n\to\infty} B_n(x) \leqq -V(-x) \leqq -L(-x) = L(x) \leqq V(x) \leqq \lim_{n\to\infty} A_n(x). \tag{2.2}$$

Now let us show that L is a Banach limit. Since B is positive, (2.2) gives that $L(x)\geqq 0$ in the case $x_k\geqq 0$ for all $k\in N$. By (2.2) it also follows that

$$\lim_{n\to\infty} B_n(x) \leqq L(x) \leqq \lim_{n\to\infty} A_n(x). \tag{2.3}$$

If we take the sequence $e=(1, 1, \dots)$ instead of (x_k) in (2.3) we easily have $L(e)=1$. Finally we have to show that $L(\sigma x)=L(x)$. If we take the sequence $\sigma x-x$ instead of x in (2.2) where $x \in m$, then, by the hypothesis, there exist strongly regular positive matrices C and D. So we get,

$$\lim_{n\to\infty} D_n(\sigma x - x) \leqq -V(x-\sigma x) \leqq L(\sigma x - x) \leqq V(\sigma x - x) \leqq \lim_{n\to\infty} C_n(\sigma x - x).$$

Therefore, since C and D are strongly regular, we have $L(\sigma x - x) = 0$, because $\lim_{n\to\infty} D_n(\sigma x - x) = \lim_{n\to\infty} C_n(\sigma x - x) = 0$ ([2], [3]). The linearity of L gives us $L(\sigma x) = L(x)$, so L is a Banach limit. Hence the functional V generates Banach limits and this completes the proof.

Throughout the paper we shall denote the set of all sublinear functionals $\Omega\colon m \to R$ which generate Banach limits by $\mathscr{B}_d$; i.e.,

$$\mathscr{B}_d = \{\Omega\colon \Omega\colon m \to R \text{ is sublinear and } \{m, \Omega\} \subset \mathscr{B}\}.$$

LEMMA 2. *Let $x \in m$. Then there exist strongly regular positive matrices A and B such that*

$$\lim_{n\to\infty} B_n(x) \leqq -\Omega(-x) \leqq \Omega(x) \leqq \lim_{n\to\infty} A_n(x) \tag{2.4}$$

for every $\Omega \in \mathscr{B}_d$.

PROOF. Let $x \in m$ and let us consider an element $\Omega \in \mathscr{B}_d$. Then $\Omega(x) \leqq W(x)$. Furthermore, there is a strongly regular positive matrix A such that $W(x) = \lim_{n\to\infty} A_n(x)$, [5]. Hence, for the matrix A, we have $\Omega(x) \leqq \lim_{n\to\infty} A_n(x)$. Since $-\Omega(-x) \geqq -W(-x)$ for $x \in m$, it can similarly be seen that there exists strongly regular positive matrix B; i.e., for the matrix B, $-\Omega(-x) \geqq \lim_{n\to\infty} B_n(x)$ is satisfied. Since $\Omega \in \mathscr{B}_d$ is arbitrary, (2.4) is satisfied for every $\Omega \in \mathscr{B}_d$.

REMARK. Let U and V be two sublinear functionals on m. If $U(x) \leqq V(x)$ for all $x \in m$ and if V generates Banach limits then U also generates Banach limits, [4]. It can easily be seen that this follows from Lemma 1.

3. A fundamental sublinear functional related to Banach limits

DEFINITION 1. We define the functional Φ as

$$\Phi(x) = \sup\{\Omega(x)\colon \Omega \in \mathscr{B}_d\}$$

where $x \in m$.

The functional $\Phi\colon m \to R$ is obviously sublinear.

Now let us consider the set $\{m, \Phi\}$. Since Φ is a sublinear functional on m, then $\{m, \Phi\} \neq \emptyset$, by Hahn—Banach extension theorem, [7, p. 121].

THEOREM 1. (i) *Φ generates Banach limits;* (ii) *Φ dominates all Banach limits.*

PROOF. (i) Let $x \in m$. By Lemma 2, there exist strongly regular positive matrices A and B such that

$$\lim_{n\to\infty} B_n(x) \leqq -\Omega(-x) \leqq \Omega(x) \leqq \lim_{n\to\infty} A_n(x) \tag{3.1}$$

for every $\Omega\in\mathscr{B}_d$. Since the inequality $\Omega(x)\leqq\lim_{n\to\infty} A_n(x)$ holds for every $\Omega\in\mathscr{B}_d$, then taking supremum over $\Omega\in\mathscr{B}_d$, we have $\Phi(x)\leqq\lim_{n\to\infty} A_n(x)$ and similarly we have $-\Phi(-x)\geqq\lim_{n\to\infty} B_n(x)$ and so, we can write the inequality

$$\lim_{n\to\infty} B_n(x) \leqq -\Phi(-x) \leqq \Phi(x) \leqq \lim_{n\to\infty} A_n(x)$$

for the strongly regular positive matrices A and B. Therefore, Φ generates Banach limits, by Lemma 1.

(ii) Since by [9], Theorem 21 (a) $W\in\mathscr{B}_d$, by the definition of Φ, we write $W(x)\leqq\Phi(x)$ for every $x\in m$. Since W dominates all Banach limits and $W(x)\leqq\leqq\Phi(x)$ for all $x\in m$, Φ also dominates all Banach limits.

THEOREM 2. *Let Q: $m\to R$ be a sublinear functional.*

(i) *Q generates Banach limits if and only if $Q(x)\leqq\Phi(x)$ for every $x\in m$.*

(ii) *Q dominates all Banach limits if and only if $\Phi(x)\leqq Q(x)$, for every $x\in m$.*

(iii) *Q both generates Banach limits and dominates all Banach limits if and only if $Q(x)=\Phi(x)$ for every $x\in m$.*

PROOF. (i) Suppose that $Q(x)\leqq\Phi(x)$ for every $x\in m$. Since Φ generates Banach limits by Theorem 1 (i), it follows immediately that Q generates Banach limits, too. Conversely, if Q generates Banach limits then, by the definition of Φ, it follows that $\Phi(x)\geqq Q(x)$ for every $x\in m$.

(ii) Now suppose that $\Phi(x)\leqq Q(x)$ for every $x\in m$. Since Φ dominates all Banach limits by Theorem 1 (ii), it follows that Q dominates all Banach limits. Conversely, if Q dominates all Banach limits then it follows by Theorem 1 (i) that $\Phi(x)\leqq Q(x)$ for every $x\in m$. Using the fact that $\{m,\Phi\}\subset\mathscr{B}\subset\{m,Q\}$, we can also reach the conclusion $\Phi(x)\leqq Q(x)$ for every $x\in m$.

(iii) This follows from (i) and (ii).

REMARK. By the conclusion obtained above, we have the following result: For $x\in m$, the sequence $x=(x_k)$ is almost convergent if and only if $\Phi(x)==-\Phi(-x)$.

4. Relations between the sublinear functionals defined on m

In this section we establish the relations between sublinear functionals defined on m and Banach limits, by using the subspace m_0 of m. First we define

$$\Psi_\Omega(x) = \inf\{\Omega(x+y):\ y\in m_0\},$$

where Ω: $m\to R$ is a sublinear functional. If $\Omega(y)\geqq 0$ for all $y\in m_0$, then Ψ_Ω is well-defined.

If Ω generates Banach limits then Ψ_Ω also generates Banach limits. But the converse is not generally true. In fact, L^* does not generate Banach limits while Ψ_{L^*} generates Banach limits, [9].

Now, let Ω be a sublinear functional which generates Banach limits. Then $L(\sigma x)=L(x)$ for every $L\in\{m, \Omega\}$ and so $\Psi_\Omega(x)=\Omega(x)$ for every $x\in m$, [4]. Hence, if we consider the sublinear functionals Φ, W and q defined by

$$q(x) = \inf_{k, n_1, n_2, \dots, n_k} \limsup_{j\to\infty} \frac{1}{k} \sum_{i=1}^{k} x_{n_i+j},$$

then we can state the following lemma.

LEMMA 3. *The following conditions are satisfied for every* $x\in m$:

(i) $\Psi_\Phi(x) = \Phi(x)$,

(ii) $\Psi_W(x) = W(x)$,

(iii) $\Psi_q(x) = q(x)$.

It was shown that q and Ψ_{L^*}, W and Ψ_S generate Banach limits and dominate all Banach limits ([4], [9], respectively). Φ also generates Banach limits and dominates all Banach limits, by Theorem 1. Therefore, we can state the following theorem by Lemma 3.

THEOREM 3. (i) $\{m, \Phi\} = \{m, \Psi_W\} = \{m, \Psi_q\} = \{m, \Psi_\Phi\} = \mathscr{B}$, *and so* (ii) $\Phi = \Psi_W = \Psi_q = \Psi_\Phi$.

5. Infinite matrices and Banach limits

Let us define the functionals $L_A^*, L_A^{**}\colon m\to R$ as follows:

$$L_A^*(x) = \limsup_{n\to\infty} \sum_{k=1}^{\infty} a_{nk} x_k$$

and

$$L_A^{**}(x) = \limsup_{n\to\infty} \sup_{j\geq 0} \sum_{k=1}^{\infty} a_{nk} x_{k+j}.$$

The functionals are sublinear and $L_A^*(x)\leqq L_A^{**}(x)$ for every $x\in m$. Furthermore, the sequence $x=(x_k)$ is A-summable if and only if $L_A^*(x)=-L_A^*(-x)$ and F_A-summable if and only if $L_A^{**}(x)=-L_A^{**}(-x)$, [9].

Now, let $A=(a_{nk})$ be a strongly regular matrix. In this case, $(A_n(x))\in c_0$ for every $x\in m_0$, [4], and so $L_A^*(y)=0$ for every $y\in m_0$ and $L_A^{**}(y)=0$. Therefore, we have

$$L_A^*(x+y) \geqq -L_A^*(-x) \quad \text{and} \quad L_A^{**}(x+y) \geqq -L_A^{**}(-x)$$

and so the functionals $\Psi_{L_A^*}$ and $\Psi_{L_A^{**}}$ are well-defined.

Now, we can give the following theorem.

THEOREM 4. *Let* $A=(a_{nk})$ *be a strongly regular matrix. Then, for every* $x\in m$, *the following equalities are satisfied:*

(i) $\Psi_{L_A^*}(x) = L_A^*(x)$,

(ii) $\Psi_{L_A^{**}}(x) = L_A^{**}(x)$.

PROOF. (i) By the definition of the functional $\Psi_{L_A^*}$, we can write

$$\Psi_{L_A^*}(x) \leqq L_A^*(x), \quad \text{for every} \quad x\in m, \tag{5.1}$$

since $(0)\in m_0$. And

$$\Psi_{L_A^*}(x) \geqq \inf \{L_A^*(x)-L_A^*(-y)\colon\ y\in m_0\}$$

since $L_A^*(x+y)\geqq L_A^*(x)-L_A^*(-y)$. Furthermore, since the matrix A is strongly regular and $y\in m_0$, we also have $-L_A^*(-y)=0$, so

$$\Psi_{L_A^*}(x) \geqq L_A^*(x), \quad \text{for every} \quad x\in m. \tag{5.2}$$

Therefore, if we compare (5.1) with (5.2), then we get

$$\Psi_{L_A^*}(x) = L_A^*(x), \quad \text{for every} \quad x\in m.$$

(ii) Let $L\in\{m, L_A^{**}\}$. Since $x\in m \Leftrightarrow \sigma x - x\in m_0\subset m$, then

$$L(\sigma x-x) \leqq L_A^{**}(\sigma x-x) \leqq \limsup_{n\to\infty} \|x\| \Big[\sum_{k=1}^{\infty} |a_{nk}-a_{n,k+1}|+|a_{n1}|\Big]$$

holds for every $x\in m$. Since A is a strongly regular matrix, the right-hand side of the above inequality is equal to zero, and so $L(\sigma x-x)\leqq 0$ for every $x\in m$. Similarly, it can be shown that $L(x-\sigma x)\leqq 0$. Therefore,

$$0 \leqq -L(x-\sigma x) = L(\sigma x-x) \leqq 0$$

is satisfied, i.e., $L(\sigma x-x)=0$. Using the linearity of L, we get $L(\sigma x)=L(x)$ and so, for every $x\in m$, $\Psi_{L_A^{**}}(x)=L_A^{**}(x)$.

If we consider this theorem and Theorem 18 in [9], we can write the following result:

COROLLARY 1. *The following conditions are equivalent:*

(i) *A is almost positive and strongly regular.*

(ii) *$\Psi_{L_A^*}$ generates Banach limits.*

(iii) *$\Psi_{L_A^{**}}$ generates Banach limits.*

6. Equality of limits

It may be remarked that the elements of $\{m, \Omega\}$, for a particular sequence, are not necessarily unique, where Ω is a sublinear functional on m. Therefore, in this section we examine the case of unique limits.

Let us consider

$$m(\Omega) = \{x \in m: \ \Omega(x) = -\Omega(-x)\},$$

where Ω is a sublinear functional on m. If we consider this definition for a special sequence $x_0 \in m$ and for all $L \in \{m, \Omega\}$ then $L(x_0)=a$, say, if and only if

$$\Omega(x_0) = -\Omega(-x_0), \tag{6.1}$$

since $L(x) \leqq \Omega(x)$ for all $L \in \{m, \Omega\}$ where $x \in m$, [3].

It is known that a bounded sequence $x=(x_k)$ is called almost convergent to a if $L(x)=a$ for every Banach limit L. The set of almost convergent sequences is denoted by F.

Let $A=(a_{nk})$ be a regular matrix. A bounded sequence (x_k) is called F_A-summable to a if

$$y_{nj} = \sum_{k=1}^{\infty} a_{nk} x_{k+j}$$

converges uniformly in j, to a, as $n \to \infty$. We denote the set of F_A-summable sequences by F_A, [6].

If we take the sublinear functionals Φ, Ψ_W, Ψ_q, Ψ_Φ instead of Ω in (6.1), then we have the following result:

COROLLARY 2. $m(\Phi)=m(\Psi_W)=m(\Psi_q)=m(\Psi_\Phi)=F$.

Proof follows easily from Theorem 3 (ii).

If we consider Theorem 4 in this paper and Theorem 6 in [3], then we have the following result:

COROLLARY 3. *Let* $A=(a_{nk})$ *be a strongly regular matrix. Then*

(i) $m(\Psi_{L_A^*}) = m \cap c_A$

(ii) $m(\Psi_{L_A^{**}}) = m \cap F_A$.

REFERENCES

[1] BANACH, S., *Théorie des opérations linéaires,* Monografie Matematyczne, Vol. 1, Warszawa, 1932. *Zbl* **5**, 209.

[2] COOKE, R. G., *Infinite matrices and sequence spaces,* Macmillan, London, 1950. *MR* **12**—694.

[3] DAS, G., Banach and other limits, *J. London Math. Soc.* (2) **7** (1974), 501—507. *MR* **49** # 924.

[4] DEVI, S. L., Banach limits and infinite matrices, *J. London Math. Soc.* (2) **12** (1975/76), 397—401. *MR* **53** # 8713.

[5] DURAN, J. P., Strongly regular matrices, almost convergence, and Banach limits, *Duke Math. J.* **39** (1972), 497—502. *MR* **46** # 9689.

[6] LORENTZ, G. G., A contribution to the theory of divergent sequences, *Acta Math.* **80** (1948), 167—190. *MR* **10**—367.

[7] Maddox, I. J., *Elements of functional analysis,* Cambridge University Press, London—New York, 1970. *MR* **52** # 11515.

[8] Petersen, G. M., *Regular matrix transformations,* McGraw-Hill, London, 1966. *MR* **37** # 642.

[9] Simons, S., Banach limits, infinite matrices and sublinear functionals, *J. Math. Anal. Appl.* **26** (1969), 640—655. *MR* 39 # 3293a.

(Received July 16, 1985)

FIRAT ÜNIVERSITESI
FEN-EDEBIYAT FAKÜLTESI
MATEMATIK BÖLÜMÜ
TR—23119 ELAZIG

ANKARA ÜNIVERSITESI
FEN-EDEBIYAT FAKÜLTESI
MATEMATIK BÖLÜMÜ
ANKARA
TURKEY

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 437—446

BIPLANES AND THEIR AUTOMORPHISMS

VLADIMIR ČEPULIĆ and MARIO ESSERT[1]

1. Introduction

Finite groups, especially permutation groups as automorphism groups of designs are a very powerful tool for investigations in design theory. Here we shall consider an application of automorphisms in constructing and characterizing biplanes. By methods, which will be explained below, we get the following result:

THEOREM. *Let $\mathcal{D}$ be a* (56, 11, 2)*-biplane admitting an automorphism of order* 8, *which fixes* 4 *points. Then $\mathcal{D}$ is isomorphic either to the Hall biplane B_{20} or to a Denniston biplane, B_{24} or B_{26}.*

Ljubo Marangunić considered in [5] the remaining case of an automorphism of order 8 fixing some point and proved that a (56, 11, 2)-biplane admitting such an automorphism is a known one.

At the beginning we shall recall some basic definitions (see [1], [2], [3]).

An incidence structure consists of a point set $\mathcal{P}$ and a line (or block) set $\mathcal{B}$, and of the incidence relation $I\subseteq\mathcal{P}\times\mathcal{B}$. Let $P\in\mathcal{P}$, $x\in\mathcal{B}$. We say that P is on x (or that x is going through P) if $(P, x)\in I$.

Denote by $\langle P\rangle=\{y\in\mathcal{B}|(P, y)\in I\}$ the set of lines through P and by $\langle x\rangle=\{Q\in\mathcal{P}|(Q, x)\in I\}$ the set of points on x. The cardinal numbers of these sets denote with $|P|$ and $|x|$, respectively.

DEFINITION 1. Let $v, k\in\mathbf{N}$ and $v>k$. A *biplane* $(v, k, 2)$ is an incidence structure $\mathcal{D}=(\mathcal{P}, \mathcal{B}, I)$ such that

(1) $|\mathcal{P}|=|\mathcal{B}|=v$,

(2) $|x|=k$ for each $x\in\mathcal{B}$,

(3) $x\neq y\Rightarrow\langle x\rangle\neq\langle y\rangle$ for all $x, y\in\mathcal{B}$,

(4) each pair of points in $\mathcal{P}$ is incident with exactly 2 common lines in $\mathcal{B}$.

In view of (3) we can consider lines as sets of points incident with them, just identifying x with $\langle x\rangle$.

Let $\mathcal{D}_1=(\mathcal{P}_1, \mathcal{B}_1, I_1)$ and $\mathcal{D}_2=(\mathcal{P}_2, \mathcal{B}_2, I_2)$ be two biplanes $(v, k, 2)$. An isomorphism from $\mathcal{D}_1$ onto $\mathcal{D}_2$ is a bijection $\alpha\colon \mathcal{P}_1\cup\mathcal{B}_1\to\mathcal{P}_2\cup\mathcal{B}_2$ such that $\mathcal{P}_1\alpha=\mathcal{P}_2$, $\mathcal{B}_1\alpha=\mathcal{B}_2$ and $(P, x)\in I_1\Leftrightarrow(P\alpha, x\alpha)\in I_2$. If $\mathcal{D}_1\equiv\mathcal{D}_2\equiv\mathcal{D}$, α is an automorphism of $\mathcal{D}$. The full automorphism group of $\mathcal{D}$ we denote with $\operatorname{Aut}\mathcal{D}$.

---

[1] The second author has done the algorithmic and programming computer work.

1980 *Mathematics Subject Classification* (1985 Revision). Primary 05B05; Secondary 51E05.
Key words and phrases. Biplane, automorphisms, chain representation.

In the following proposition we resume some known properties of biplanes, which we shall use frequently.

PROPOSITION 1. *An incidence structure is a biplane* $(v, k, 2)$, $v, k \in \mathbf{N}$, *if and only if*

(1) $|\mathscr{P}| = |\mathscr{B}| = v = k(k-1)/2+1$,

(2) $|x| = |P| = k, \quad k > 2$ *for all* $P \in \mathscr{P}, \quad x \in \mathscr{B}$,

(3) $|\langle x\rangle \cap \langle y\rangle| = |\langle P\rangle \cup \langle Q\rangle| = 2$ *for all* $P, Q \in \mathscr{P}, \quad x, y \in \mathscr{B}$.

The condition $|\langle x\rangle \cap \langle y\rangle| = 2$ in (3) we shall call the *consistence condition.*

Although the number of biplanes might be infinite, until now there are up to isomorphism only 17 known biplanes, one of each $(4, 3, 2)$, $(7, 4, 2)$, $(11, 5, 2)$, three $(16, 6, 2)$, four $(37, 9, 2)$, five $(56, 11, 2)$ and two $(79, 13, 2)$, the recent one $(56, 11, 2)$ being discovered in 1985 by Z. Janko and Tran van Trung (see [4]) after about 7 years of stagnancy.

Denoting points with $0, 1, \ldots, v-1$ and lines as sets of points, the smallest 3 biplanes are as follows ($\bar{0}$ stands for 10):

(4, 3, 2)	(7, 4, 2)		(11, 5, 2)		
0 1 2	0 1 2 3	1 2 5 6	0 1 2 3 4	0 4 7 9 $\bar{0}$	2 3 6 7 9
0 1 3	0 1 4 5	1 3 4 6	0 1 5 6 7	1 2 7 8 $\bar{0}$	2 4 5 6 $\bar{0}$
0 2 3	0 2 4 6	2 3 4 5	0 2 5 8 9	1 3 5 9 $\bar{0}$	3 4 5 7 8
1 2 3	0 3 5 6		0 3 6 8 $\bar{0}$	1 4 6 8 9	

2. Chain representation of a biplane in a given basis $\langle P\rangle$

Here we introduce another representation of a biplane $\mathscr{D} = (\mathscr{P}, \mathscr{B}, I)$ (see also [6]). Let $P \in \mathscr{P}$ be a chosen point, which we call the basic point. Denote with $0, 1, \ldots, k-1$ the lines through P, thus $\langle P\rangle = \{0, 1, \ldots, k-1\}$. By Prop. 1 (3). for any $Q \in \mathscr{P} \setminus \{P\}$ it holds $\langle P\rangle \cap \langle Q\rangle = \{a(Q), b(Q)\}$ and then also $\langle a(Q)\rangle \cap \langle b(Q)\rangle = \{P, Q\}$. Thus there is a one-to-one correspondence between points $Q \in \mathscr{P} \setminus \{P\}$ and unordered pairs of lines in $\langle P\rangle$. In the following we shall identify $Q \equiv a(Q)b(Q) \equiv ab$, $a, b \in \langle P\rangle$. The representation of points just introduced leads to a new representation of lines in $\mathscr{B} \setminus \langle P\rangle$ as well.

Let $x \in \mathscr{B} \setminus \langle P\rangle$ and $Q_1 \equiv a_1 a_2 \in \langle x\rangle$ with $a_1, a_2 \in \langle P\rangle$. Now $\langle a_2\rangle \cap \langle x\rangle = \{Q_1, Q_2\}$ by Proposition 1 (3) with $Q_2 \equiv a_2 a_3$ and again $\langle a_3\rangle \cap \langle x\rangle = \{Q_2, Q_3\}$. Here $a_3 \neq a_1$ since otherwise $Q_1 \equiv Q_2$. If $Q_3 = Q_1$ then $\langle Q_1\rangle \cap \langle Q_2\rangle \supseteq \{x, a_2, a_3\}$, a contradiction. Thus $Q_1 \neq Q_3$ also.

Assume that we have already built a sequence $Q_1 \equiv a_1 a_2, \ldots, Q_{t-1} = a_{t-1} a_t$, with $Q_i \in \langle x\rangle$ and $a_i \neq a_j$ for $i \neq j$. Let $\langle a_t\rangle \cap \langle x\rangle = \{Q_{t-1}, Q_t\}$ and $Q_t \equiv a_t a_{t+1}$. If $a_{t+1} = a_i$ for some i, $1 < i < t$, then $\langle a_i\rangle \cap \langle x\rangle \supseteq \{Q_{i-1}, Q_i, Q_t\}$, a contradiction. Thus either $a_{t+1} = a_1$ or $a_{t+1} \neq a_i$ for all i, $1 < i < t$. Because of finiteness of $\langle P\rangle$ there must be some t such that $a_{t+1} = a_1$ and we come in this way to a closed sequence $a_1 a_2, a_2 a_3, \ldots, a_{t-1} a_t, a_t a_1$. For brevity we write it as $/a_1 a_2 \ldots a_t/$, where each pair of neighbour symbols denotes a point, the both symbols adjacent to delimiters being also considered as neighbours. The symbol $/a_1 \dots a_t/$ just introduced is called a chain of length t.

If there exist other points on x not included in the above chain, we proceed in the same manner and build further chains. Thus, whole x will be represented by one or more chains. Such a representation we call a chain representation of x in the basis $\langle P \rangle$ and identify it with x. For chain representation it holds

PROPOSITION 2. *Let* $\mathscr{D}=(\mathscr{P}, \mathscr{B}, I)$ *be a biplane* $(v, k, 2)$, $P \in \mathscr{P}$,

$$\langle P \rangle = \{0, 1, \dots, k-1\}$$

and $Q \in \mathscr{P} \setminus \{P\}$, $x \in \mathscr{B} \setminus \langle P \rangle$. *Then* $\langle P \rangle \cap \langle Q \rangle = \{a, b\}$ *is uniquely determined by* Q, *and we identify* $Q \equiv ab$. *The line* x *is represented by chains in the form*

$$(*) \qquad x = /a_1^{(1)} \dots a_{t_1}^{(1)}/ \dots /a_1^{(s)} \dots a_{t_s}^{(s)}/$$

where the chains satisfy the following conditions:

(1) *each pair of neighbour symbols denotes a point in the above sense, the two symbols adjacent to delimiters inside a chain being also considered as neighbours;*
(2) *each symbol of* $\langle P \rangle$ *occurs in* x *just once;*
(3) *each chain is of length at least* 3;
(4) *two different lines have exactly two pairs of neighbour symbols in common;*
(5) *any succesive three neighbour symbols of a line* x *occur as such neighbours only in* x.

PROOF. The assertions (1)—(4) were proved at the beginning of this paragraph or follow immediately from what was stated there. (5): If $\dots abc \dots$ belongs to a chain of x and also to a chain of $y \in \mathscr{B} \setminus \langle P \rangle$, then $\langle ab \rangle \cap \langle bc \rangle \supseteqq \{x, y, b\}$, which contradicts Proposition 1 (3). Thus (5) is also proved.

As the chains themselves are unoriented and cyclic, it will be useful to introduce a canonical form of chain representation. We shall do it by the following

DEFINITION 2. We say that a chain representation of a line x

$$(*) \qquad x \equiv /a_1^{(1)} \dots a_{t_1}^{(1)}/ \dots /a_1^{(j)} \dots a_{t_j}^{(j)}/ \dots /a_1^{(s)} \dots a_{t_s}^{(s)}/$$

is *canonical* if and only if

(1) $a_2^{(j)} = \min \{a_i^{(j)};\ 1 \leqq i \leqq t_j\}$ for each j, $1 \leqq j \leqq s$,

(2) $a_1^{(j)} < a_3^{(j)}$ for each j, $1 \leqq j \leqq s$,

(3) $a_2^{(j')} < a_2^{(j'')}$ for $j' < j''$.

If $(*)$ is canonical, we say that x is of *type* $t_1 - t_2 - \dots - t_s$ in the basis $\langle P \rangle$.

Obviously, the canonical form of a line x in the given basis $\langle P \rangle$ is unique, and $a_2^{(1)} = 0$ for this form.

EXAMPLES. The line /5241/6987/03$\bar{0}$/ (here again $\bar{0}$ stands for 10) has the canonical form /30$\bar{0}$/4152/7698/ and has chain type 3—4—4. For $k=11$ there are 13 chain types: 3—4—4, 4—3—4, 4—4—3, 3—3—5, 3—5—3, 5—3—3, 5—6, 6—5, 4—7, 7—4, 3—8, 8—3, 11.

By Proposition 2 (5) each canonical line x is uniquely determined within a biplane by its beginning triplet $a_1^{(1)}0a_3^{(1)}=a0b$. We say that x belongs to the *layer* $a0b$ and denote $x=a0b\dots$.

3. Representation of automorphisms fixing some point

We shall consider now the action of automorphisms fixing some point of a biplane represented in the corresponding basis.

DEFINITION 3. Let $\mathscr{D}=(\mathscr{P},\mathscr{B},I)$ be a biplane $(v,k,2)$ and $\varrho\in\operatorname{Aut}\mathscr{D}$. We denote $f_\varrho(\mathscr{P})=\{Q\in\mathscr{P}|Q\varrho=Q\}$, $f_\varrho(\mathscr{B})=\{y\in\mathscr{B}|y\varrho=y\}$ and $F_\varrho=|f_\varrho(\mathscr{P})|$, $f_\varrho=|f_\varrho(\mathscr{B})|$. Henceforth $\mathscr{D}$ will always denote $\mathscr{D}=(\mathscr{P},\mathscr{B},I)$, a biplane $(v,k,2)$. The next proposition is of basic importance.

PROPOSITION 3. *Let $P\in\mathscr{P}$ and let $G\leqq(\operatorname{Aut}\mathscr{D})_{\mathscr{P}}$ be a subgroup of the stabilizer of P in $\operatorname{Aut}\mathscr{D}$, i.e. $PG=P$. Then $\langle P\rangle G=\langle P\rangle$ also, and G acts faithfully on $\langle P\rangle$ as a permutation group. If $\varrho\in G$ and $Q\equiv ab\in\mathscr{P}\setminus\{P\}$, with $a,b\in\langle P\rangle$, then $Q\varrho\equiv\equiv(a\varrho)(b\varrho)$. Similarly, if $x\in\mathscr{B}\setminus\langle P\rangle$ is represented in $\langle P\rangle$, we get the $\langle P\rangle$-representation of $x\varrho$ by replacing symbols from $\langle P\rangle$ with their ϱ-images.*

PROOF. Obviously, $\langle P\rangle G=\langle P\rangle$. If G does not act faithfully on $\langle P\rangle$, then there must be a $\varrho\in G$, $\varrho\neq 1$, which acts trivially on $\langle P\rangle$. Now, for any $Q\equiv ab\in\mathscr{P}\setminus\langle P\rangle$, we have $\langle P\varrho\rangle\cap\langle Q\varrho\rangle=\langle P\rangle\cap\langle Q\varrho\rangle=\{a\varrho,b\varrho\}$, and thus $Q\varrho\equiv(a\varrho)(b\varrho)\equiv ab\equiv Q$ implying $f_\varrho(\mathscr{P})=\mathscr{P}$ and $f_\varrho(\mathscr{B})=\mathscr{B}$. The last assertion follows from the fact that $\langle x\rangle\varrho=\{Q\varrho|Q\in\langle x\rangle\}$.

Now we prove the following

PROPOSITION 4. *Let $P\in\mathscr{P}$, and $\varrho\in(\operatorname{Aut}\mathscr{D})_{\mathscr{P}}$, $|\varrho|=p^f$, p a prime. If α_s is the number of cycles of length p^s, $0\leqq s\leqq f$, in the representation of ϱ on $\langle P\rangle$, then the number β_s of ϱ-orbits in $\mathscr{P}$ of length p^s is equal:*

1) *Case* $p>2$.

$$\beta_0=\binom{\alpha_0}{2}+1,$$

$$\beta_s=(\sum_{j=0}^{s-1}\alpha_j p^j+((\alpha_s-1)/2)p^s+(p^s-1)/2)\alpha_s,\quad \textit{for}\quad s\geqq 1.$$

2) *Case* $p=2$.

$$\beta_0=\binom{\alpha_0}{2}+\alpha_1+1,$$

$$\beta_s=(\sum_{j=0}^{s-1}\alpha_j 2^j+((\alpha_s-1)/2)2^s+(2^s-2)/2)\alpha_s+\alpha_{s+1},\quad \textit{for}\quad s\geqq 1.$$

PROOF. 1) $p>2$.

We have $Q\equiv ab\equiv Q\varrho\equiv(a\varrho)(b\varrho)$ if and only if $a\varrho=a$, $b\varrho=b$. There are exactly $\binom{\alpha_0}{2}$ such pairs. As P is also a fixed point, β_0 is as asserted. Every symbol of a p^j-cycle of ϱ in $\langle P\rangle$, $j<s$, combined with the first symbol of any p^s-cycle as a point, determines a ϱ-orbit of length p^s in $\mathscr{P}$ and all such orbits are different. There are $\alpha_j p^j \alpha_s$ such orbits. Similarly, combining each symbol of a p^s-cycle with the first symbol of another p^s-cycle we get different p^s-orbits again. There are $\binom{\alpha_s}{2} p^s=((\alpha_s(\alpha_s-1))/2)p^s$ such orbits. Finally, every pair of two symbols in a p^s-cycle determines a point of a p^s-orbit. Thus there are $\alpha_s \binom{p^s}{2}: p^s=\alpha_s(p^s-1)/2$ such orbits.

2) $p=2$.

The proof is analogous as above with exception of the case, where both symbols belong to the same 2^s-cycle. Let $(a_1\dots a_{2^s})$ be such a cycle and $Q\equiv a_i a_j$. If $Q\varrho^t\equiv Q$, $t\geqq 1$, then $(a_i\varrho^t)(a_j\varrho^t)\equiv a_i a_j$. Here $t<2^s$ if and only if $a_i\varrho^t=a_j$ $a_j\varrho^t=a_i$, i.e. $a_i\varrho^{2t}=a_i$. This implies $t=2^{s-1}$ and thus $a_j=a_i\varrho^{2^{s-1}}$. There are exactly 2^{s-1} such pairs $a_1 a_{2^{s-1}+1}, \dots, a_{2^{s-1}} a_{2^s}$ in our cycle, which belong to a 2^{s-1} orbit. Thus we have $\left[\binom{2^s}{2}-2^{s-1}\right]\alpha_s$: $2^s=((2^s-2)/2)\alpha_s$ ϱ-orbits of length 2^s arising from 2^s-cycles and α_{s+1} ϱ-orbits of length 2^s from 2^{s+1}-cycles.

As we are especially interested in the case, when ϱ fixes an incident pair $(P, x)\in I$, we shall consider this case separately.

PROPOSITION 5. *Let $\varrho\in \mathrm{Aut}\,\mathscr{D}$ and let there exist such $P\in\mathscr{P}$, $0\in\langle P\rangle$, that $P\varrho=P$, $0\varrho=0$. Then the point orbits of ϱ have the same distribution as the line orbits of ϱ, i.e. for every s there are as many point-orbits in $\mathscr{P}$ of length s as line-orbits in $\mathscr{B}$ of the same length.*

PROOF. Let $\langle P\rangle=\{0, 1, \dots, k-1\}$ and let ϱ be represented as permutation on $\langle P\rangle$. We shall construct a bijection α of $\mathscr{P}$ onto $\mathscr{B}$, α: $\mathscr{P}\to\mathscr{B}$, such that $Q_1, Q_2\in\mathscr{P}$ belong to the same ϱ-orbit if and only if $Q_1\alpha, Q_2\alpha\in\mathscr{B}$ belong to the same orbit. Let $Q\in\mathscr{P}\setminus\{P\}$ and $Q\equiv ab$, with $a, b\in\langle P\rangle$. Consider first the case $a\neq 0$, $b\neq 0$. We set $Q\alpha=a0b\ \dots$, the line in $\mathscr{D}$ belonging to the layer $a0b$.

Now

$$(Q\varrho^t)\alpha = [(ab)\varrho^t]\alpha = [(a\varrho^t)(b\varrho^t)]\alpha = (a\varrho^t)0(b\varrho^t)\dots = (a\varrho^t)(0\varrho^t)(b\varrho^t)\dots =$$
$$= (a0b\dots)\varrho^t = (Q\alpha)\varrho^t$$

by Proposition 3. By Proposition 2 (5) we see that $Q\varrho^t\neq Q$ if and only if $(Q\alpha)\varrho^t\neq$ $\neq Q\alpha$. Consider next the case $a=0$, $b\neq 0$. We set $Q\alpha=b$. Now $(Q\varrho^t)\alpha=[(0b)\varrho^t]\alpha=$ $=[0(b\varrho^t)]\alpha=b\varrho^t=(Q\alpha)\varrho^t$. Again, $Q\varrho^t\neq Q$ if and only if $(Q\alpha)\varrho^t\neq Q\alpha$. Set finally $P\alpha=0$. The bijection α maps obviously s-orbits of $\mathscr{P}$ onto s-orbits of $\mathscr{B}$, proving our proposition.

4. Combinatorial construction of biplanes admitting an automorphism fixing an incident pair

The general idea of constructing biplanes is the following: to generate biplanes line by line (layer by layer). Each new line must be consistent with all the previous lines.

We shall sketch here the outlines of an algorithm for constructing biplanes in chain form. In the following we identify biplane $\mathscr{D}$ with the set of its lines, considered as sets of points. We assume that there is an automorphism group G acting on $\mathscr{D}$, i.e. $\mathscr{D}G=\mathscr{D}$. Moreover, we suppose that there is a G-stable incident pair $(P, 0)\in I$, so that $PG=P, 0G=0$. Let G and $\mathscr{D}$ be represented in the basis $\langle P\rangle$ and $N=N_{\Sigma_{\langle P\rangle}}(G)$, $\Sigma_{\langle P\rangle}$ being the symmetric group on $\langle P\rangle$. For $\alpha\in\Sigma_{\langle P\rangle}$, it is obviously $\mathscr{D}\alpha$ isomorphic to $\mathscr{D}$, where α acts on lines of $\mathscr{D}$ as explained in Proposition 3.

In the rest of this chapter we always suppose that the above assumptions are satisfied and we use the above notation.

PROPOSITION 6. *Let $\alpha\in N$. Then $\mathscr{D}\alpha$ is also a biplane admitting G as automorphism group.*

PROOF. We have $(\mathscr{D}\alpha)G=\mathscr{D}(\alpha G)=\mathscr{D}(G\alpha)=(\mathscr{D}G)\alpha=\mathscr{D}\alpha$, which proves the assertion.

DEFINITION 4. Each layer $a0b$ of $\mathscr{D}$ generates under G a layer orbit $(a0b)G=\{(a\gamma)0(b\gamma)|\gamma\in G\}$ of length $|G:H_{ab}|$, where $H_{ab}=\{\gamma\in G|\{a, b\}\gamma=\{a, b\}\}$. For each such layer G-orbit we choose a representative, briefly: *layer representative*, and define an ordering among them. Let $\mathscr{D}'$ be a set of full G-orbits of lines in the defined order of layer representatives such that each two lines of $\mathscr{D}'$ satisfy the consistence condition. Such a line set $\mathscr{D}'$ will be called a *beginning scheme*.

PROPOSITION 7. *Let $\mathscr{D}'$ be a beginning scheme, which is not yet a biplane. Let x be a line belonging to a layer representative $a0b$, which is outside of $\mathscr{D}'$, such that $\mathscr{D}''=\mathscr{D}'\cup xG$ is again a beginning scheme. Then the following holds:*

(1) It must be $xH_{ab}=x$, i.e. $x\gamma=x$ for all $\gamma\in H_{ab}$. This is equivalent with $xH_{ab}^*=x$ for any minimal generator set H_{ab}^* of H_{ab}.

(2) If x is consistent with all y in $\mathscr{D}'$, then for each $\gamma\in G$, $x\gamma$ is also consistent with all y in $\mathscr{D}'$.

(3) Let $G_{ab}^*, G_{ab}^*\subseteq G$, be a set of minimal cardinality satisfying the condition: $G_{ab}^\#=\{\gamma|\gamma\in G_{ab}^*$ or $\gamma^{-1}\in G_{ab}^*\}$ is a set of right coset representatives for $(G/H_{ab})^\#=\{H_{ab}\gamma|\gamma\in G\setminus H_{ab}\}$. Then the consistence conditions $|\langle x\rangle\cap\langle x\rangle\gamma|=2$ for all $\gamma\in G$ are equivalent with $|\langle x\rangle\cap\langle x\rangle\gamma|=2$ for all $\gamma\in G_{ab}^*$. Such a set G_{ab}^* will be called a minimal consistence set for the group G and the layer $a0b$.

PROOF. (1) If x belongs to the layer $a0b$ and $\gamma\in H_{ab}\subseteq G$, then $x\gamma$ is in $\mathscr{D}''$ and $x\gamma$ belongs to the layer $a0b$, too. By Proposition 2 (5) it must be $x\gamma=x$. The second part of the assertion is trivial.

(2) It is $|\langle x\rangle\gamma\cap\langle y\rangle|=|\langle x\rangle\gamma\gamma^{-1}\cap\langle y\rangle\gamma^{-1}|=|\langle x\rangle\cap\langle y\rangle\gamma^{-1}|$. But $y\gamma^{-1}\in\mathscr{D}'$ if $y\in\mathscr{D}'$, by the definition of $\mathscr{D}'$. Thus $|\langle x\rangle\cap\langle y\rangle\gamma^{-1}|=2$ by our assumption and therefore also $|\langle x\rangle\gamma\cap\langle y\rangle|=2$.

(3) The assertion follows immediately from $|\langle x\rangle\gamma\cap\langle x\rangle|=|\langle x\rangle\cap\langle x\rangle\gamma^{-1}|$.

Next we define the precedence relation for lines and beginning schemes.

DEFINITION 5. Let $\mathcal{D}_1'$ and $\mathcal{D}_2'$ be two beginning schemes of the same length (i.e. containing the same layer orbits), represented in the canonical chain form. We assume that there is a fixed ordering defined among chain types. Let $x_1 \in \mathcal{D}_1'$, $x_2 \in \mathcal{D}_2'$ be two lines belonging to the same layer. We say that x_1 *precedes* x_2, $x_1 \leqq x_2$, if

either: (1) the chain type of x_1 precedes that of x_2,
or: (2) x_1 and x_2 have the same chain type and the symbol sequence of x_1 precedes that of x_2 lexicographically.

We say that $\mathcal{D}_1'$ precedes $\mathcal{D}_2'$, $\mathcal{D}_1' \leqq \mathcal{D}_2'$, if $\mathcal{D}_1'$ precedes $\mathcal{D}_2'$ lexicographically in terms of their lines belonging to the layer representatives and written in the defined order (see Definition 4). As usual, $<$ will denote $\leqq$ and $\neq$, in both cases.

In our seeking for biplanes, we try to omit the isomorphic ones, retaining only those, which are first in the order of the defined precedence. Thus, in view of Proposition 6, if $\mathcal{D}'$ is a beginning scheme and $\mathcal{D}'\alpha < \mathcal{D}'$, we omit such $\mathcal{D}'$.

Now, we can explain our algorithm for constructing all biplanes $(v, k, 2)$ admitting some automorphism group G.

Let S be any ordered set, and $s \in S$. We denote with s^+ the next element in S, and with $(S)_0$ the first element of S. If s is the last element of S, then $s^+ \notin S$.

Let $M \equiv M_{ab}$ be the set of all combinatorially possible lines, belonging to the layer representative $a0b$, generated in the order defined by Definition 5. Let H_{ab}^* be a minimal generator set of H_{ab} and G_{ab}^* a minimal consistence set, both as ordered sets (in arbitrary way). Then we can build the *layer set* $\Lambda \equiv \Lambda_{ab}$ consisting of all lines in M, which are consistent with their G-images by the following

Λ-Procedure

```
x=(M)_0;
repeat
  γ:=(H*_ab)_0;
  while γ∈H*_ab and xγ=x do γ:=γ+;
  if γ∉H*_ab then
  {γ=(G*_ab)_0;
    while γ∈G*_ab and |⟨x⟩∩⟨x⟩γ|=2 do γ:=γ+;
    if γ∉G*_ab then write x to Λ;
  };
  x:=x+;
until x+∉M.
```

After building the layer sets we can go over to the main

Algorithm

Let $\triangle_i$, $i=1, \ldots, t$ denote the ordered set of the beginning schemes built up to the i-th layer representative $a_i 0 b_i$ and Λ_i the layer set of $a_i 0 b_i$. We set $\triangle_0 = \{\emptyset\}$. The elements of the sets $\triangle_i$, Λ_{i+1}, N will be denoted by $\mathcal{D}'$, x and α, respectively.

The following procedure leads to the construction of all biplanes, which satisfy the above conditions:

$\triangle_0 = \{\emptyset\}$; $i := 0$;
repeat
  $\triangle_{i+1} = \emptyset$; $\mathscr{D}' = (\Delta_i)_0$;
  repeat
    $x := (\Lambda_{i+1})_0$;
    repeat
      if x is consistent with $\mathscr{D}'$ then
        $\{\mathscr{D}'' := \mathscr{D}' \cup \{x\gamma | \gamma \in G\}$; $\alpha := (N)_0$;
        while not $(\mathscr{D}''\alpha < \mathscr{D}'')$ and $\alpha^+ \in N$ do $\alpha := \alpha^+$;
        if not $(\mathscr{D}''\alpha < \mathscr{D}'')$ then write $\mathscr{D}''$ to $\triangle_{i+1}$;
        $\}$;
      $x := x^+$;
    until $x \notin \Lambda_{i+1}$;
    $\mathscr{D}' := \mathscr{D}'^+$;
  until $\mathscr{D}' \notin \triangle_i$;
  $i := i + 1$;
until $i = t$ or $\triangle_i = \emptyset$;
if $\triangle_i = \emptyset$ then there is none solution
else $\triangle_t$ is the set of solutions.

The above algorithm is a relatively powerful tool for finding out all biplanes (56, 11, 2) with given automorphism groups of order >3, fixing some incident pair, even with personal computers. We get our results just in this way.

5. Proof of the Theorem

We turn now over to the case of biplanes (56, 11, 2) admitting an automorphism group $G = \langle \varrho \rangle$ of order 8, which fixes 4 points. Let P be a point fixed by ϱ. Then we can represent ϱ as a permutation on $\langle P \rangle = \{0, 1, 2, \ldots, \bar{0} \equiv 10\}$ of the form $\varrho = 012(3456789\bar{0})$. Here $\alpha_0 = 3$, $\alpha_1 = 0$, $\alpha_2 = 0$, $\alpha_3 = 1$ and by Proposition 4 (2) we have $\beta_0 = 4$, $\beta_1 = 0$, $\beta_2 = 1$, $\beta_3 = 6$. Thus there are four 1-orbits, one 4-orbit and six 8-orbits for ϱ on $\mathscr{P}$ and on $\mathscr{B}$. (The only other possibility for an automorphism of order 8 fixing some point would be $\varrho = 0(12)(3456789\bar{0})$ with 2 fixed points and 2 fixed lines.) In $\langle P \rangle$ there are 3 fixed lines and one 8-orbit. Consequently, in $\mathscr{B} \setminus \langle P \rangle$ there are: one 1-orbit, one 4-orbit and five 8-orbits. As one easily sees, the corresponding layer representatives are: 102 for 1-orbit, 307 for 4-orbit, and 103, 203, 304, 305, 306 for 8-orbits. We shall build biplanes just in this order of layer representatives. The chain type order will be the same as in the example of Chapter 2.

It is $N = N_{\Sigma\langle P \rangle}(\langle \varrho \rangle) = \langle \alpha, \beta, \varrho, \gamma, \delta \rangle$ with $\alpha = (012)$, $\beta = (12)$, $\gamma = (46)(59)(8\bar{0})$ and $\delta = (4\bar{0})(59)(68)$. Here $\varrho^\gamma = \varrho^3$ and $\varrho^\delta = \varrho^7$. Furthermore, we can obviously choose the following sets as H^*_{ab} and G^*_{ab}:

$$H^*_{12} = \{\varrho\}, \quad G^*_{12} = \emptyset; \quad H^*_{37} = \{\varrho^4\}, \quad G^*_{37} = \{\varrho, \varrho^2\};$$

$$H^*_{13} = H^*_{23} = H^*_{34} = H^*_{35} = H^*_{36} = \emptyset,$$

$$G^*_{13} = G^*_{23} = G^*_{34} = G^*_{35} = G^*_{36} = \{\varrho, \varrho^2, \varrho^3, \varrho^4\}.$$

By simple combinatorial calculations based on canonical form in Definition 7, one can easily check that there are 67259 combinatorially possible lines in each layer. It can be easily proved that $A_1=\{x_1=102/5397/64\bar{0}8/,\ x_2\equiv 102/43\bar{0}98765/,\ x_3\equiv$ $\equiv 102/6385\bar{0}749/\}$. Since $x_3\gamma=x_2$, it is $\triangle_1=\{x_1, x_2\}$. The same can be reached by machinal computation. Applying the algorithm of Chapter 4 with all its reductions, we get with help of computer as the only solutions the following biplanes (we write only the ϱ-orbits representatives, the other lines one gets acting on these representatives with the permutation ϱ):

1)
1 0 2 / 5 3 9 7 / 6 4 10 8 /
3 0 7 / 5 1 9 2 / 6 4 8 10 /
1 0 3 / 4 2 6 7 / 8 5 9 10 /
2 0 3 / 8 1 10 7 / 5 4 6 9 /
3 0 4 / 5 1 8 6 / 7 2 10 9 /
3 0 5 / 2 1 6 10 / 7 4 9 8 /
3 0 6 / 9 1 10 4 / 7 2 8 5 /

2)
1 0 2 / 5 3 9 7 / 6 4 10 8 /
3 0 7 / 5 1 9 2 / 6 4 8 10 /
1 0 3 / 4 2 6 / 9 5 10 7 8 /
2 0 3 / 4 1 6 7 / 8 5 9 10 /
3 0 4 7 2 10 5 / 8 1 9 6 /
3 0 5 4 9 7 8 / 2 1 6 10 /
3 0 6 7 2 8 9 / 5 1 10 4 /

3)
1 0 2 / 4 3 10 9 8 7 6 5 /
3 0 7 / 4 1 8 2 / 6 5 9 10 /
1 0 3 / 9 2 10 7 / 6 4 8 5 /
2 0 3 / 6 1 9 10 5 7 4 8 /
3 0 4 / 5 1 7 10 / 6 2 8 9 /
3 0 5 / 9 1 10 8 / 4 2 7 6 /
3 0 6 / 2 1 5 9 / 7 4 10 8 /

4)
1 0 2 / 4 3 10 9 8 7 6 5 /
3 0 7 / 4 1 8 2 / 6 5 9 10 /
1 0 3 / 5 2 6 9 10 8 4 7 /
2 0 3 / 5 1 10 7 9 4 8 6 /
3 0 4 5 10 1 8 7 2 9 6 /
3 0 5 7 1 6 8 9 2 4 10 /
3 0 6 4 10 7 8 / 2 1 9 5 /

Representing these biplanes by other basis, changing the basic point, one can easily see (see Appendix in [6]) that the first is isomorphic to Hall's biplane B_{20}, the second to Denniston's biplane B_{21} and the last two to Denniston's biplane B_{26}.

This proves the assertion of our theorem.

ACKNOWLEDGEMENT. We are thankful to professor Zvonimir Janko from Mathematisches Institut Heidelberg, for the suggestion of this theme.

REFERENCES

[1] Beutelspacher, A., *Einführung in die endliche Geometrie*, I, Blockpläne, Bibliographisches Institut, Mannheim, 1982. *MR* **84j**: 05015.

[2] Dembowski, P., *Finite Geometries*, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 44, Springer, Berlin—New York, 1968. *MR* **38** # 1597.

[3] Hughes, D. R. and Piper, F. C., *Design theory*, Cambridge University Press, Cambridge, 1985. *MR* **87e**: 05022.

[4] Janko, Z. and Trung, T. Van, A new biplane of order 9 with a small automorphism group, *J. Combin. Theory Ser. A* **42** (1986), 305—309. *MR* **87j**: 51024.

[5] Maranguníć, L., The classification of biplanes (56, 11, 2) admitting an automorphism of order 8 which fixes some point, *Rad. Mat.* **2** (1986), 99—111. *MR* **87h**: 51016.

[6] Salwach, C. J. and Mezzaroba, J. A., The four known biplanes with $k=11$, *Internat. J. Math. Sci.* **2** (1979), 251—260. *MR* **80g**: 05014a.

(Received September 20, 1985)

ELEKTROTEHNIČKI FAKULTET
UNSKA 3
YU—41000 ZAGREB
CROATIA
YUGOSLAVIA

FAKULTET STROJARSTVA I BRODOGRADNJE
ĐURE SALAJA 5
YU—41000 ZAGREB
YUGOSLAVIA

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 447—452

COMPLETE BASES IN TOPOLOGICAL SPACES II

JOSE L. BLASCO

Introduction

The word space will refer to Tychonoff spaces. A base[1] $\mathscr{D}$ on a space X is called complete if it coincides with the trace on X of all zero-sets in its Wallman realcompactification $\upsilon(X, \mathscr{D})$. The existence of noncomplete bases is not trivial. In the matter the author proved in ([3], Theorem 2) the following sufficient condition:

(C) *Let K be a compactification of a space X such that there exist two compact subsets C_1 and C_2 of $K \sim \mathbf{Q}(X, K)$ which are disjoint and homeomorphic. If $C_1 \cup C_2$ is not z-separated from X, then there exist noncomplete bases on X.*

This condition was then applied to obtain the following results:

A paracompact space X is Lindelöf if and only if each base on X is complete ([3], Corollary 4).

If G is a topological group, then υG is Lindelöf if and only if each base on G is complete ([4], Theorem 4).

In this paper we prove that condition (C) is also necessary for the existence of noncomplete bases. In addition, we give two equivalent conditions to (C) which are more convenient to apply. From those conditions we prove that for a separable space X, υX is Lindelöf if and only if each base on X is complete. In the same way we characterize the Lindelöf property of certain products of realcompact spaces.

As an application of our results we give a partial solution to a question on certain inverse-closed subalgebras of $C(X)$ posed by Hager and Johnson in [6].

Preliminaries

Inverse images of one-point sets under a mapping φ are called fibers of φ. As usual, $C(X)$ will denote the algebra of all continuous real-valued functions on a space X. The set of points of X where a member f of $C(X)$ is equal to zero is called the zero-set of f and will be denoted by $Z(f)$. The collection of all zero-sets of X

[1] This notion is due to E. F. Steiner [8] who uses the term separating nest generated intersection ring. An equivalent concept is the strong delta normal base due to R. A. Alò and H. L. Shapiro [1].

1980 *Mathematics Subject Classification* (1985 Revision). Primary 54D20; Secondary 54D35.

Key words and phrases. Nest generated intersection ring, strong delta normal base, complete base, Lindelöf space, inverse-closed subalgebra of $C(X)$, algebra on X, realcompact space.

will be denoted by $Z(X)$. The $\mathbf{Q}$-closure of a subset A of X is the set $\mathbf{Q}(A, X)$ of all points $p \in X$ for which each G_δ-set containing p meets A.

Compactifications will always be Hausdorff. We write βX for the Stone—Čech compactification of the space X. It is well-known that $\upsilon X = \mathbf{Q}(X, \beta X)$. Two extensions T_1 and T_2 of X are said to be equivalent if they are homeomorphic via a map that leaves X pointwise fixed. In this case we write $T_1 = T_2$.

Let E and F be two subsets of the space X. The set F is said to be z-separated from E if there is a zero set Z in X such that $F \subset Z$ and $Z \cap E = \emptyset$. It is known [7] that *X is a Lindelöf space if and only if there is a compactification K of X such that every compact subset of $K \sim X$ is z-separated from X.*

LEMMA 1. *Let h be a continuous mapping from a space X into a space Y, let F be a subset of X and let M be a closed subset of Y which misses $h(F)$. Then:*

(a) *If M is z-separated from $h(F)$, then $h^{-1}(M)$ is z-separated from F.*

(b) *If X is compact and $h^{-1}(M)$ is z-separated from F, then M is z-separated from $h(F)$.*

PROOF. (a) Is immediate. (b) Let Z be a zero-set in X such that $h^{-1}(M) \subset Z$ and $Z \cap F = \emptyset$. Then $X \sim Z = \bigcup \{Z_n \colon n \in N\} \supset F$, $Z_n \in Z(X)$, $n \in N$, and therefore $M \cap (\bigcup \{h(Z_n) \colon n \in N\}) = \emptyset$. Since M is closed and $h(Z_n)$ is compact, it follows that M is z-separated from $h(Z_n)$ for every $n \in N$. Consequently, M is z-separated from $\bigcup \{h(Z_n) \colon n \in N\}$.

LEMMA 2. *Let K be a compactification of a space X. Let $\mathscr{E}$ be an upper semi-continuous decomposition of a compact subset M of $K \sim \mathbf{Q}(X, K)$ whose members are compact and are z-separated from X.*

If $\mathscr{E}_0 = \mathscr{E} \cup \{\{p\} \colon p \in K \sim M\}$, then the decomposition space $K^* = K(\mathscr{E}_0)$ is a compactification of X such that $\mathbf{Q}(X, K) = \mathbf{Q}(X, K^*)$.

PROOF. Clearly K^* is a compact space containing X as a dense subspace. Since $\mathscr{E}_0$ is an upper semicontinuous decomposition of K into compact sets, then K^* is a (Hausdorff) compactification of X. If φ is the quotient map from K onto K^*, it is easy to check that $\varphi(\mathbf{Q}(X, K)) \subset \mathbf{Q}(X, K^*)$. Now let p be a point in $K \sim \mathbf{Q}(X, K)$. Since the members of $\mathscr{E}$ are z-separated from X it follows that $\varphi^{-1}(\varphi(p))$ is z-separated from X. According to Lemma 1 (b) the point $\varphi(p)$ is z-separated from X (in K^*), therefore $\varphi(p) \notin \mathbf{Q}(X, K^*)$ and $\mathbf{Q}(X, K^*) \subset \varphi(\mathbf{Q}(X, K))$. Since the restriction of φ to X is the identity we have $\mathbf{Q}(X, K^*) = \mathbf{Q}(X, K)$.

LEMMA 3. *Let K be a compactification of a space X and let $\mathscr{F}$ be a base for a filter of closed sets in X. If $\mathscr{F}$ is closed under countable intersections and $\bigcap \{F \colon F \in \mathscr{F}\} = \emptyset$, then $K_0 = \bigcap \{\mathrm{cl}_K F \colon F \in \mathscr{F}\}$ is a compact subset of $K \sim X$ which is not z-separated from X.*

PROOF. It is clear that K_0 is a non-empty compact subset of $K \sim X$. Let Z be a zero-set in K such that $K_0 \subset Z$. Every compact subset of $K \sim Z$ is disjoint from some $\mathrm{cl}_K F$, $F \in \mathscr{F}$. Since $K \sim Z$ is σ-compact, Z must contain the intersection of countably many $\mathrm{cl}_K F$. Since the intersection of these sets F is not empty, necessarily $Z \cap X \neq \emptyset$.

The results

We will write $\omega(X, \mathscr{D})$ (resp. $\upsilon(X, \mathscr{D})$) for the Wallman compactification (resp. Wallman realcompactification) associated with a given base $\mathscr{D}$ on a space X. For definitions and basic results the reader is referred to [1], [2] and [9]. For a base $\mathscr{D}$ on X let $\hat{\mathscr{D}}$ be the trace on X of all zero-sets in $\upsilon(X, \mathscr{D})$. Then $\hat{\mathscr{D}}$ is a base on X which contains $\mathscr{D}$. A base $\mathscr{D}$ on X is called complete if $\mathscr{D}=\hat{\mathscr{D}}$. It follows from ([2], Corollaries 2.1 and 2.2) that:

(i) *A base $\mathscr{D}$ on X is complete if and only if $\beta(\upsilon(X, \mathscr{D}))=\omega(X, \mathscr{D})$.*
(ii) *$\hat{\mathscr{D}}$ is the largest base on X such that $\upsilon(X, \hat{\mathscr{D}})=\upsilon(X, \mathscr{D})$.*
(iii) *$\hat{\mathscr{D}}$ is the smallest complete base on X containing $\mathscr{D}$.*

THEOREM 4. *For every space X, the following statements are equivalent:*
(1) *In some compactification K of X there exist disjoint compact subsets K_1 and K_2 of $K\sim X$ such that:* (a) *$K_1\cup K_2$ is not z-separated from X.* (b) *For each $i=1, 2$ there is a continuous map φ_i from K_i onto a compact space C whose fibers are z-separated from X.*
(2) *Condition* (1) *with $K=\beta X$.*
(3) *There is a compactification K^* of X with two compact subsets C_1 and C_2 of $K^*\sim\mathbf{Q}(X, K^*)$ which are disjoint, homeomorphic and whose union is not z-separated from X.*
(4) *There exist noncomplete bases on X.*

PROOF. (1) *implies* (2). Let ψ be the Stone extension of the identity from X into K and put $K_i'=\psi^{-1}(K_i)$, $i=1, 2$. Thus K_1' and K_2' are disjoint compact subsets of $\beta X\sim X$, and according Lemma 1 (b) the set $K_1'\cup K_2'$ is not z-separated from X. Finally, it is clear that $\varphi_i\circ\psi$ is a continuous mapping from K_i' onto C whose fibers are z-separated from X, $i=1, 2$.

(2) *implies* (3). Firstly, since the fibers of φ_1 and φ_2 are z-separated from X it follows that $(K_1\cup K_2)\cap\mathbf{Q}(X, \beta X)=\emptyset$. The family $\mathscr{E}=\{\varphi_1^{-1}(x), \varphi_2^{-1}(x)\colon x\in C\}$ is an upper semicontinuous decomposition of $K_1\cup K_2$ into compact sets which are z-separated from X. By Lemma 2 we have that $\mathbf{Q}(X, K^*)=\mathbf{Q}(X, \beta X)$ where $K^*=K(\mathscr{E}_0)$. Moreover if ψ is the Stone extension of the identity from X into K^*, it is clear that $\psi(K_1)$ and $\psi(K_2)$ are disjoint compact subsets of $K^*\sim\mathbf{Q}(X, K^*)$ which are homeomorphic (to C). By Lemma 1 (a), the set $\psi(K_1)\cup\psi(K_2)$ is not z-separated from X.

(3) *implies* (4) follows from Theorem 2 in [3].

(4) *implies* (1). Let $\mathscr{D}$ be a noncomplete base on X. Let S be a subset of X. For convenience we write $\mathrm{cl}_\omega S$ (resp. $\mathrm{cl}_\upsilon S$, $\mathrm{cl}_\beta S$) for the closure of S in $\omega(X, \mathscr{D})$ (resp. $\upsilon(X, \mathscr{D})$, $\beta(\upsilon(X, \mathscr{D}))$). By ([2], Corollary 2.1) $\beta(\upsilon(X, \mathscr{D}))\neq\omega(X, \mathscr{D})$ and therefore there exist disjoint zero-sets Z_1' and Z_2' in $\upsilon(X, \mathscr{D})$ such that $K=\mathrm{cl}_\omega Z_1'\cap\cap\mathrm{cl}_\omega Z_2'$ is a non-empty compact subset of $\omega(X, \mathscr{D})\sim\upsilon(X, \mathscr{D})$. Put $Z_i=Z_i'\cap X$.

Firstly, let us see that if D_1 and D_2 are members of $\mathscr{D}$ such that $Z_i\subset D_i$, then $D_1\cap D_2\neq\emptyset$. Suppose that $D_1\cap D_2=\emptyset$, then $\mathrm{cl}_\omega D_1\cap\mathrm{cl}_\omega D_2=\emptyset$ and hence $\mathrm{cl}_\omega Z_1\cap\cap\mathrm{cl}_\omega Z_2=\emptyset$. Since each zero-set in $\upsilon(X, \mathscr{D})$ meets X, we have that $\mathrm{cl}_\upsilon Z_i=Z_i'$ and hence $\mathrm{cl}_\omega Z_i = \mathrm{cl}_\omega Z'_i$, $i=1,2$. Then $\mathrm{cl}_\omega Z'_1 \cap \mathrm{cl}_\omega Z'_2 = \emptyset$ which is a contradiction. Consequently, $D_1 \cap D_2 \neq \emptyset$.

Consider the family

$$\mathscr{F} = \{D \in \mathscr{D}\colon\ D_1 \cap D_2 \subset D,\ Z_i \subset D_i,\ D_i \in \mathscr{D},\ i = 1, 2\}$$

and let us see that $\mathscr{F}$ is a $\mathscr{D}$-filter closed under countable intersections. Given a sequence $\{D_n\colon n \in N\}$ in $\mathscr{F}$, we have $D_1^n \cap D_2^n \subset D_n$, $Z_i \subset D_i^n$, $D_i^n \in \mathscr{D}$, $i=1, 2$, $n \in N$. Then

$$Z_i \subset \bigcap \{D_i^n\colon\ n \in N\} = D'_i \in \mathscr{D}, \quad i = 1, 2,$$

and since Z_1 and Z_2 are not separated by members of $\mathscr{D}$, it follows that

$$\emptyset \neq D'_1 \cap D'_2 \subset \bigcap \{D_n\colon\ n \in N\} = D$$

therefore $D \in \mathscr{F}$.

Since $\{\mathrm{cl}_\omega D\colon D \in \mathscr{D}\}$ is a base for the closed sets in $\omega(X, \mathscr{D})$ such that $\mathrm{cl}_\omega (D_1 \cap D_2) = \mathrm{cl}_\omega D_1 \cap \mathrm{cl}_\omega D_2$ for $D_1, D_2 \in \mathscr{D}$ it is easy to check that

$$(*) \qquad \emptyset \neq K = \bigcap \{\mathrm{cl}_\omega D\colon\ D \in \mathscr{F}\} \subset \omega(X, \mathscr{D}) \sim \upsilon(X, \mathscr{D}).$$

If ψ is the Stone extension of the identity from $\upsilon(X, \mathscr{D})$ into $\omega(X, \mathscr{D})$ we have that $\psi^{-1}(K)$ is a non-empty compact subset of $\beta(\upsilon(X, \mathscr{D})) \sim \upsilon(X, \mathscr{D})$. Put $K_0 = \bigcap \{\mathrm{cl}_\beta D\colon D \in \mathscr{F}\}$ and let us see that $\psi(K_0) = K$.

Since ψ is continuous it follows that $\psi(K_0) \subset K$ and hence $K_0 \subset \psi^{-1}(K)$. Now consider a point z in K. The family

$$\mathscr{U}_z = \{D \in \mathscr{D}\colon\ z \in \mathrm{cl}_\omega D\}$$

is a $\mathscr{D}$-ultrafilter and therefore a base for a $\hat{\mathscr{D}}$-filter. If p is a point in

$$\bigcap \{\mathrm{cl}_\beta D\colon\ D \in \mathscr{U}_z\}$$

then $\psi(p) \in \mathrm{cl}_\omega F$ for every $F \in \mathscr{U}_z$.

Since

$$\{z\} = \bigcap \{\mathrm{cl}_\omega F\colon\ F \in \mathscr{U}_z\}$$

it follows that $\psi(p) = z$. As z is in K, from $(*)$ we have that $\mathscr{F} \subset \mathscr{U}_z$ and hence $p \in K_0$. Thus $\psi(K_0) = K$.

Applying Lemma 3 to $\beta(\upsilon(X, \mathscr{D}))$ and $\mathscr{F}$ it follows that K_0 is not z-separated from X.

Define $K_i = \psi^{-1}(K) \cap \mathrm{cl}_\beta Z_i$. Then $\psi(K_i) \subset K \cap \mathrm{cl}_\omega Z_i = K$ and since $\psi(\mathrm{cl}_\beta Z_i) = \mathrm{cl}_\omega Z_i$, it follows that $\psi(K_i) = K$, $i = 1, 2$.

Now consider the following cases:

(i) There exists $F \in \mathscr{F}$ such that $F \cap Z_1 = \emptyset$. Then $\mathrm{cl}_\beta F \cap \mathrm{cl}_\beta Z_1 = \emptyset$ and therefore $K_0 \cap K_1 = \emptyset$. Moreover, $\psi(K_0) = \psi(K_1) = K$ and since $K \cap \upsilon(X, \mathscr{D}) = \emptyset$, it follows that the fibers of ψ in K_0 and K_1 are z-separated from X.

(ii) For every $F \in \mathscr{F}$, $F \cap Z_1 \neq \emptyset$ and $F \cap Z_2 \neq \emptyset$. It is easy to see that

$$K_0 \cap \mathrm{cl}_\beta Z_1 = \bigcap \{\mathrm{cl}_\beta (F \cap Z_1)\colon\ F \in \mathscr{F}\}.$$

Since $\mathscr{V} = \{F \cap Z_1\colon F \in \mathscr{F}\}$ is a base for a $\hat{\mathscr{D}}$-filter and $\mathscr{V}$ is closed under countable intersections, from Lemma 3 it follows that $K_0 \cap \mathrm{cl}_\beta Z_1$ is a compact subset of K_1 which is not z-separated from X. Then $K_1 \cap K_2 = \emptyset$, $\psi(K_1) = \psi(K_2) = K$ and K_1 is not z-separated from X. As in (i) the fibers of ψ in K_1 and K_2 are z-separated from X.

By an algebra on a space X is meant a subring of $C(X)$ which contains the constants, separates points and closed sets, and is closed under uniform convergence and inversion in $C(X)$. If A is an algebra on X and $Z(A) = \{Z(f): f \in A\}$, the map $A \to Z(A)$ is a one-to-one correspondence between the family of all algebras on X and the family of all bases on X ([9], Theorem 4.3). In [6], Hager and Johnson have proved that if the Hewitt realcompactification υX is Lindelöf then each algebra on X is isomorphic to $C(Y)$ for some space Y and they ask whether the converse of this assertion holds. From ([3], Theorem 10) and ([4], Theorem 4) the answer is yes if X is paracompact or a topological group.

COROLLARY 5. *Let X be a separable space. The following conditions are equivalent:*

(1) *υX is Lindelöf.*
(2) *Each base on X is complete.*
(3) *Each algebra on X is a $C(Y)$.*

PROOF. (1) *implies* (3) follows from ([6], 4.4). (3) *implies* (2) is an immediate consequence of ([2], Theorem 5). (2) *implies* (1). Suppose that υX is not Lindelöf and let K be a compact subset of $\beta X \sim \upsilon X$ which is not z-separated from X. Since each point in $\beta X \sim \upsilon X$ is z-separated from υX, then K is infinite. Let $p_1, p_2 \in K$, $p_1 \neq p_2$ and choose zero-sets in K, K_1 and K_1', such that $K = K_1 \cup K_1'$, $p_1 \in K_1 \sim K_1'$ and $p_2 \in K_1' \sim K_1$. Then one of these sets, say K_1, is not z-separated from X. Since $p_2 \notin \upsilon X$ there is a zero-set $Z \in Z(\beta X)$ such that $p_2 \in Z$ and $Z \cap (K_1 \cup \upsilon X) = \emptyset$. By ([5], 9.11) Z contains a copy M of βN.

On the other hand, since X is separable there is a continuous map σ from βN onto βX. Put $K_2 = \sigma^{-1}(K_1)$ and consider K_2 as a subset of the copy M of βN. Then there is a continuous map from the compact subset K_2 of M onto K_1, whose fibers are z-separated from X because $M \subset Z$. From Theorem 4 with $C = K_1$, there exist noncomplete bases on X.

COROLLARY 6. *Let X and Y be realcompact spaces having more than one point. If each base on $X \times Y$ is complete then $X \times Y$ is Lindelöf.*

PROOF. Firstly, assume that $\beta X \sim X$ has at most one point. Then X is pseudocompact ([5], 6J) and therefore compact. By ([3], Proposition 6) Y is Lindelöf, hence $X \times Y$ is Lindelöf and the result is proved.

Now suppose that $X \times Y$ is not Lindelöf. Then $\beta X \sim X$ has at least two points and there is a compact subset K' of $(\beta X \times \beta Y) \sim (X \times Y)$ which is not z-separated from $X \times Y$. We shall prove that there exist a point $z \in \beta X \sim X$ and a compact subset K_1 of K' such that $z \notin p_X(K_1)$ and K_1 is not z-separated from $X \times Y$.

Let u and v be distinct points of $\beta X \sim X$ and denote the sets $p_X^{-1}(\{u\}) \cap K'$ and $p_X^{-1}(\{v\}) \cap K'$ by K_u' and K_v', respectively. If one of these sets, say K_u', is empty, put $z = u$ and $K_1 = K'$. Assume that K_u' and K_v' are non-empty. If $C_1 = K_u'$ and $C_2 = K_v'$, then C_1 and C_2 are disjoint compact subsets of K', therefore there is a function $f \in C(\beta X \times \beta Y)$ such that $f(C_1) = \{0\}$ and $f(C_2) = \{1\}$. If $E = \{(x, y) \in K': f(x, y) \leq 3/4\}$ and $F = \{(x, y) \in K': f(x, y) \geq 1/4\}$, then one of these sets, say E, is not z-separated from $X \times Y$. In this case put $K_1 = E$ and $z = v$.

Define $K_2=\{z\}\times p_Y(K_1)$. Then K_2 is a compact subset of $(\beta X\times\beta Y)\sim(X\times Y)$ disjoint from K_1 and the map $\varphi\colon K_1\to K_2$ defined by $\varphi(x,y)=(z,y)$ is continuous and onto. We will prove that $\varphi^{-1}(z,y)$ is z-separated from $X\times Y$ for each $y\in p_Y(K_1)$.

Let y be a point in $Y\cap p_Y(K_1)$. Since $K_1\cap(X\times Y)=\emptyset$, it follows that $K_1^y=$ $=p_X(p_Y^{-1}(\{y\})\cap K_1)$ is a compact subset of $\beta X\sim X$. According to ([3], Proposition 6) X is Lindelöf, therefore there is a function $g\in C(\beta X)$ such that $K_1^y\subset Z(g)$ and $Z(g)\cap X=\emptyset$. The function h defined by $h(x,t)=g(x)$ is continuous on $\beta X\times\beta Y$,

$$\varphi^{-1}(z,y)=K_1^y\times\{y\}\subset Z(h)$$

and $Z(h)\cap(X\times Y)=\emptyset$.

Now let $y\in p_Y(K_1)\sim Y$. Since Y is realcompact there is a function $w\in C(\beta Y)$ such that $w(y)=0$ and $Z(w)\cap Y=\emptyset$. The function h defined by $h(x,t)=w(t)$ is continuous on $\beta X\times\beta Y$, $\varphi^{-1}(z,y)\subset Z(h)$ and $Z(h)\cap(X\times Y)=\emptyset$. According to Theorem 4 there exist noncomplete bases on $X\times Y$.

COROLLARY 7. *Let $\{X_i\colon i\in I\}$ be a family of realcompact spaces. Assume that at least two X_i's have more than one point. Then the following conditions are equivalent:*

(1) *The product space $X=P\{X_i\colon i\in I\}$ is Lindelöf.*
(2) *Each base on X is complete.*
(3) *Each algebra on X is a $C(Y)$.*

PROOF. (1) *implies* (3) follows from ([6], 4.4). (3) *implies* (2) is a consequence of ([2], Theorem 5). (2) *implies* (1) is a consequence of Corollary 6 because the conditions on the X_i's imply that X can be factored as a product of two subproducts with more than one point both of them.

REFERENCES

[1] ALÒ, R. A. and SHAPIRO, H. L., *Normal topological spaces*, Cambridge Tracts in Mathematics, No. 65, Cambridge University Press, New York, 1974. *MR* **52** # 11808.
[2] BLASCO, J. L., Complete bases and Wallman realcompactifications, *Proc. Amer. Math. Soc.* **75** (1979), 114—118. *MR* **80g**: 54033.
[3] BLASCO, J. L., Complete bases in topological spaces, *Studia Sci. Math. Hungar.* **20** (1985), 49—54.
[4] BLASCO, J. L., A note on inverse-closed subalgebras of $C(X)$, *Acta Math. Hungar.* (to appear).
[5] GILLMAN, L. and JERISON, M., *Rings of continuous functions*, Van Nostrand Co., Princeton, 1960. *MR* **22** # 6994.
[6] HAGER, A. W. and JOHNSON, D. G., A note on certain subalgebras of $C(X)$, *Canad. J. Math.* **20** (1968), 389—393. *MR* **36** # 5697.
[7] SMIRNOV, YU. M., On normally disposed sets of normal spaces, *Mat. Sbornik (N.S.)* **29** (1951), 173—176 (in Russian). *MR* **13**—371.
[8] STEINER, E. F., Normal families and completely regular spaces, *Duke Math. J.* **33** (1966), 743—745. *MR* **33** # 7975.
[9] STEINER, A. K. and STEINER, E. F., Nest generated intersection rings in Tychonoff spaces, *Trans. Amer. Math. Soc.* **148** (1970), 589—601. *MR* **41** # 7637.

(Received November 5, 1986)

FACULTAD DE CIENCIAS MATEMATICAS
DR MOLINER 50
BURJASOT
VALENCIA
SPAIN

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 453—460

REMARKS ON THE OPTIMIZATION OF THE LEHMER—SCHUR METHOD

A. GALÁNTAI

Abstract

Optimal or near optimal Lehmer—Schur methods utilizing a maximum of seven disks are determined. A lower estimation for the optimal parameters is given for the general case. This estimation is found to be sharper than Friedli's estimation (3).

1. Introduction

Lehmer's method was introduced in 1961 [5] for the approximate solution of polynomial equations of the form

$$p(z) = a_n z^n + a_{n-1} z^{n-1} + \ldots + a_1 z + a_0 = 0 \qquad (a_i \in \mathbf{C}). \tag{1.1}$$

This method is based on a geometrical search technique and a Schur-criterion which may be used to decide whether or not a polynomial has a zero within a given disk of the complex plane.

The Lehmer—Schur method is defined as follows

(i) Choose an initial disk which includes at least one zero of the polynomial.

(ii) Cover the inclusion disk by a finite set of disks of smaller radii.

(iii) Choose a disk from the covering such that it includes a zero of the polynomial.

(iv) Go to (ii).

If the coverings of step (ii) are similar to a fixed covering of the unit disk the algorithm is obviously convergent with linear speed.

Lehmer's choice consists of 9 disks (see (5), (6) and (8) pp. 391, P. 50). The first disk has radius $1/2$ and is concentric with the unit disk. The remaining 8 disks have radius $2/5$.

The computational efficiency of Lehmer's method which obviously depends on the choice of covering was investigated by Henrici (4) and Friedli (3). In fact, Henrici investigated the more general class of proximity methods to which Lehmer's method belongs. Using some results from discrete geometry Henrici found an optimal method in the class of congruent coverings of the unit disk. (A covering is called congruent if its disks are congruent.) Henrici's covering consists of 8 congruent disks of radius 0.44504... . Friedli approached the problem in a different way and proved that for 3 or more disks there exists an optimal Lehmer—Schur method and it is defined by an extremal q-covering of the unit disk. (A covering is

1980 *Mathematics Subject Classification.* Primary 65H05; Secondary 52A45.
Key words and phrases. Polynomial equations, optimal covering algorithms.

called q-covering if the radii of its disks form a geometric progression with ratio q.) Friedli gave no extremal q-covering for any number of disks. However, using a heuristic technique he found two excellent q-coverings which consist of 11 and 22 disks, respectively. Friedli also gave a lower estimation for the parameter of the extremal q-coverings.

In this paper we give a sharper lower estimation for these parameters and construct the extremal q-coverings (optimal Lehmer—Schur methods) for 3 and 4 disks, respectively. Near optimal coverings are constructed also for the case of 5, 6 and 7 disks. An important consequence of our estimation is that the parameter of Friedli's 22 disk q-covering approaches the asymptotic optimum with an error not exceeding 0.02. Finally, we investigate the effect of changing the measure of comparison. It is shown that the optimal methods will change only slightly and Henrici's method will not be optimal in the subclass of congruent coverings. The optimal Lehmer—Schur method in the class of congruent coverings is also determined.

2. Preliminaries

Denote by $C^{(l)}=\{D_1, D_2, \ldots, D_l\}$ $(l\geqq 3)$ a covering of the unit disk D_0 $(D_0\subset \mathbf{C})$ such that each disk D_i satisfies $r_i<1$ $(i=1, 2, \ldots, l)$ where $r_i=r(D_i)$ denotes the radius of disk D_i. We assume that the coverings of step (ii) are similar to a fixed covering $C^{(l)}$. The disks of the covering can be tested in arbitrary but fixed order and the numbering of disks corresponds to this fixed order. If the zero testing on one disk is chosen as the unit of computational cost and all disks are tested then the computational efficiency index of the Lehmer—Schur method is defined by

$$\varphi(C^{(l)}) = \max\{i/|\log r_i|:\ i = 1, 2, \ldots, l\}, \tag{2.1}$$

where $C^{(l)}$ denotes the covering related to the method ([3]). A method (covering $\hat{C}^{(l)}$) is said to be optimal (extremal) if

$$\varphi(\hat{C}^{(l)}) = \inf\{\varphi(C^{(l)}):\ C^{(l)},\ l \text{ fixed}\}. \tag{2.2}$$

The choice of the base of the logarithm does not affect the optimal method. If we choose the base of the logarithm to be 10 as in [3] then Lehmer's method has efficiency index 22.616 and Henrici's method has 22.753.

A covering $C^{(l)}$ is said to be a q-covering of the unit disk if there is a number $0<q<1$ such that $r_i=q^i$ $(i=1, 2, \ldots, l)$. For q-coverings (denoted by $C_q^{(l)}$) the efficiency index becomes $\varphi(C_q^{(l)})=1/|\log q|$. If $Q(l)$ denotes the extremal q-covering of the unit disk by l disks, that is $Q(l)$ is the smallest q value for which a q-covering of the unit disk by l disks exists, then it can be proven ([3]) that

$$\inf\{\varphi(C^{(l)}):\ C^{(l)},\ l \text{ fixed}\} = \varphi(C_{Q(l)}^{(l)}). \tag{2.3}$$

Hence it is sufficient to seek only extremal q-coverings. Friedli's heuristic constructions have the parameters $l=11$, $q=0.7698$ and $l=22$, $q=0.7663$, respectively. His lower estimation for $Q(l)$ is the positive zero of the equation $\sum_{i=1}^{l} q^{2i}=1$. Since $Q(l)$ is monotonically decreasing as $l\to+\infty$ the best efficiency index is approached asymptotically and the limit is

$$\lim_{l\to+\infty} (1/|\log Q(l)|). \tag{2.4}$$

Friedli's estimation yields the lower bounds $\lim_{l\to+\infty} Q(l) \geqq 1/\sqrt{2}$ for the smallest q-parameter and $\lim_{n\to+\infty} (1/|\log Q(l)|) \geqq 6.643$ for the best efficiency index.

The efficiency index (2.1) is related to the maximum computational cost in a given class of polynomial equations. It is also possible to introduce an average computational efficiency index which is based on assumptions from geometrical probability ([3], [4]). For Lehmer's method this average efficiency index is 11.143 and for Henrici's choice it is 11.168. Since for q-coverings this measure coincides with (2.1) ([3]) we do not investigate it separately.

3. The lower estimation and the optimal constructions

Consider the following equation

$$\sum_{i=1}^{l} \arcsin q^i = \pi \qquad (l \geqq 3). \tag{3.1}$$

This equation has a unique real solution in the interval $(0, 1)$. This unique solution is denoted by $q(l)$. Then we have

THEOREM 1. *For the optimal parameter $Q(l)$ of the extremal q-coverings the inequality*

$$Q(l) \geqq q(l) \qquad (l \geqq 3) \tag{3.2}$$

holds. If $l=3$ or $l=4$ then equality holds in (3.2) *and the coverings of Figure 1 and 2 give the extremal (optimal) configurations (Lehmer—Schur methods).*

PROOF. First we prove that there is no q-covering $C_q^{(l)}$ of the unit disk such that $q<q(l)$ is satisfied. If there was such a q-covering then its disks must cover the unit circle, too. The arc defined by the intersection of the disk of radius r_i $(r_i<1)$ and the unit circle has a maximum length given by $2 \arcsin r_i$ (see Fig. 3). Thus the unit circle can be covered by the disks of $C_q^{(l)}$ only if the inequality

$$2 \sum_{i=1}^{l} \arcsin r_i = 2 \sum_{i=1}^{l} \arcsin q^i \geqq 2\pi \tag{3.3}$$

is satisfied. Since $q(l)$ denotes the unique solution of (3.1) in $(0, 1)$ the inequality (3.3) cannot be satisfied for any number $q<q(l)$. From this fact (3.2) clearly follows. For $q=q(l)$ the disks of radii q^i $(i=1, 2, \ldots, l)$ can cover the unit circle if they are arranged as in Figure 3. However, they do not necessarily cover the whole unit disk. From Molnár's [7] more general result it follows that these disks also cover the unit disk in the case of 3 and 4 disks (see Figure 1 and 2). More precisely

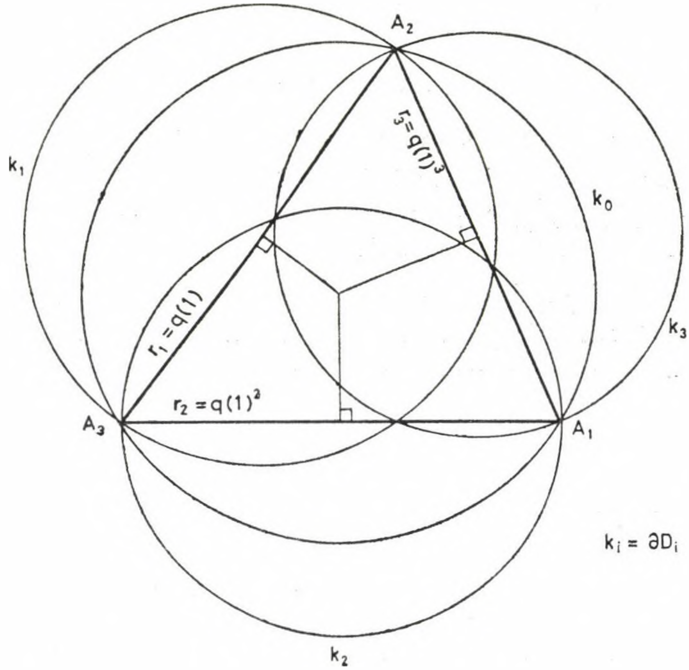

Fig. 1

the disks cover the triangle $A_1A_2A_3$ or the quadrangle $A_1A_2A_3A_4$ inscribed in the unit circle. Hence they cover the whole unit disk. Thus the configurations of Figure 1 and 2 are the extremal q-coverings and the related Lehmer—Schur methods are optimal for $l=3, 4$.

It can be shown that $q(l)<Q(l)$ for $l\geqq 5$. However, utilizing concepts from Theorem 1 we can construct near optimal q-coverings for $l=5, 6, 7$. Let us arrange the disks of radii $r_i=q(l)^i$ $(i=1, 2, \dots, l)$ such that they cover the unit circle in the following (counterclockwise) order

$$1, 2\left[\frac{l-1}{2}\right]+1, \dots, 5, 3, 2, 4, 6, \dots, 2\left[\frac{l}{2}\right]. \tag{3.4}$$

Let the intersection of disk D_i and the unit circle be denoted by the arc A_iB_i. Increase the parameter $q(l)$ by a small number $\delta>0$ and the radii of the disks to $r_i'=(q(l)+\delta)^i$ $(i=1, 2, \dots, l)$. Then move the center of the disk of radius r_i' toward the center of the unit circle such that the arc defined by their intersection exactly coincides with the arc A_iB_i $(i=1, 2, \dots, l)$. If this new configuration is a covering of the unit disk the procedure is finished. If not we increase the value of δ and repeat the process.

The following table includes some of the results we obtained in this way as well as some of the results obtained by Friedli.

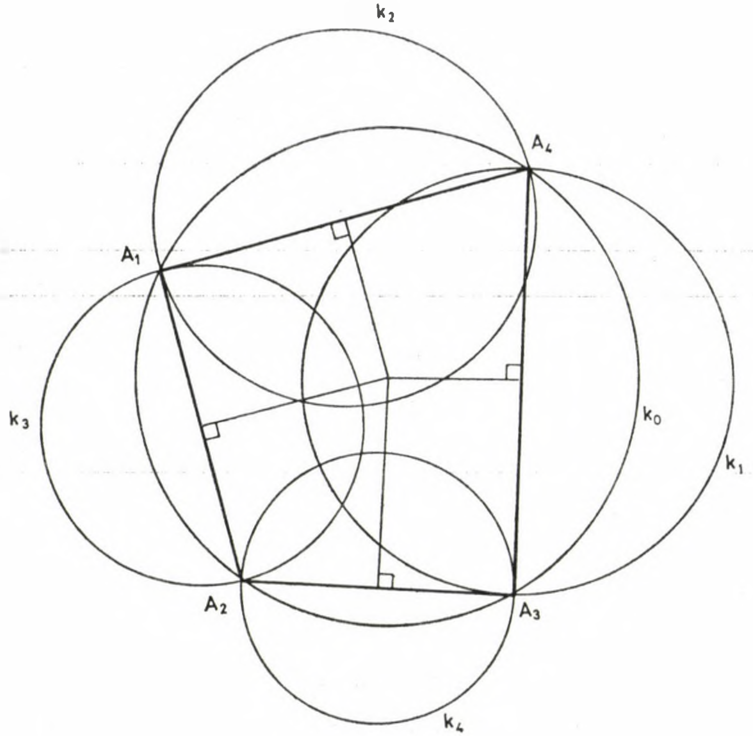

Fig. 2

Table 1

	Friedli's lower estimation	$q(l)$	The q-parameter of the best known construction
3	0.73735	0.926968	0.926968
4	.72027	.860545	.860545
5	.71320	.820865	.8210
6	.71003	.796856	.7992
7	.70853	.781736	.7895
8	.70780	.771866	.7840
$+\infty$	$\sqrt{2}/2$	0.748987*	0.7663**

The number marked by $(*)$ is an approximation to the value $\lim_{l\to+\infty} q(l)$ with error less than 0.0003. The other number marked by $(**)$ is the parameter of Friedli's configuration with 22 disks.

The efficiency index of the 4-disk optimum is 15.332 which is 67.8% of the computational efficiency of Lehmer's method. The coverings we gave for $l=5, 6, 7$ are near optimal in the sense that none of the disks can be omitted from the covering of the unit circle and their q-parameters approximate $Q(l)$ with an absolute error not exceeding 0.008. If we seek a q-covering for which the (average) efficiency index is less than the average computational efficiency of Lehmer's method and the number of disks is minimal we need at least 6 disks. For our 6-disk configuration the (average) efficiency index is 10.273 which is less than 11.143. This configuration is given in Table 2 where c_i denotes the center of disk D_i of radius r_i.

Table 2

i	$\operatorname{Re} c_i$	$\operatorname{Im} c_i$	r_i
1	0.543001636	0	0.7992
2	−0.691665726	−0.128510216	0.63872064
3	−0.406039690	0.683577178	0.510465536
4	−0.239074846	−0.818753509	0.407964057
5	0.281818842	0.845796974	0.326044874
6	0.348926122	−0.849297389	0.260575064

By observing Table 1, we find that our lower estimation is obviously sharper than Friedli's result. It follows from our estimation that his 22-disk configuration approaches $\lim_{l\to+\infty} Q(l)$ with an error not exceeding 0.02.

4. Remarks

The computational efficiency index (2.1) is based on the assumption that all the disks are tested in (algorithm) step (ii). However, it is clear that we need to test only the first $l-1$ disks. In this case the computational efficiency index changes to

$$\tilde{\varphi}(C^{(l)}) = \max\{i/|\log r_i|,\ (l-1)/|\log r_l|:\ i = 1, \dots, l-1\}. \tag{4.1}$$

For Lehmer's covering the $\tilde{\varphi}$-efficiency index is 20.104 and for Henrici's covering it is 19.91. Having adopted the technique of [3] we can introduce the concept of quasi q-coverings. A covering $C^{(l)}$ is said to be a quasi q-covering if the radii of disks satisfy

$$r_i = q^i \quad (i = 1, 2, \dots, l-1), \quad r_l = q^{l-1} \quad (0 < q < 1).$$

We can then prove that the $\tilde{\varphi}$-optimal methods are based on the extremal quasi q-coverings with minimal q-parameter. We can also reformulate all of the results of Section 3 in a slightly different but similar form.

It is worth noting that Henrici's covering is not $\tilde{\varphi}$-optimal in the subclass of congruent coverings. For a congruent covering $C^{(l)}$ with common radius r the efficiency index is $\tilde{\varphi}(C^{(l)}) = (l-1)/|\log r|$. If $R(l)$ denotes the minimal radius for which a congruent covering of the unit disk by l disks still exists, then

$$\inf\{\tilde{\varphi}(C^{(l)}):\ C^{(l)} \text{ congruent, } l \text{ fixed}\} = (l-1)/|\log R(l)|. \tag{4.2}$$

Obviously, the best method based on congruent coverings is defined by $\inf\{(l-1)/|\log R(l)|:\ l \geqq 3\}$. The extremal value $R(l)$ and the related extremal con-

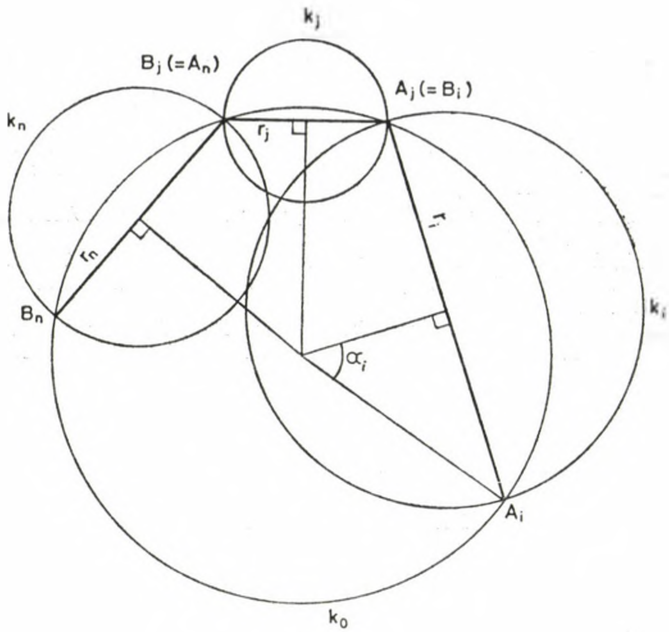

Fig. 3

figuration are known only for a few values of l ([1], [9]). For $l=3, 4, 7$ the extremal radii are $R(3)=\sqrt{3}/2$, $R(4)=\sqrt{2}/2$, $R(7)=1/2$ and the extremal configurations can be found in [9]. For $l=5, 6$ only a conjecture of Grünbaum and Neville was known ([9]), but upon correcting these conjectures, Bezdek [1] determined the extremal configurations. The extremal radii are given by $R(5)=0.6098$ and $R(6)=0.5559$ with four decimal precision. No other result is known. However, one can prove the trivial inequality $R(l)>1/\sqrt{l}$. For the range $7<l<12$ we need to prove the sharper inequality

$$R(l) > \sin(\pi/(l-1)) \qquad (l > 7). \tag{4.3}$$

Assume that there exists an l-disk congruent covering of the unit disk with radius $\sin(\pi/(l-1))$. Note that $\sin(\pi/(l-1))<1/2$ for $l>7$. Hence there must be a disk (say disk l) which has no common point with the unit circle. The first $l-1$ disks of radius $\sin(\pi/(l-1))$ can cover the unit circle only in one configuration which is similar to Figure 3. The diameter of the uncovered area of the unit disk is $2(1-2\sin^2(\pi/(l-1)))$ which is greater than 1. Since the last disk cannot cover this area we have proved (4.3). Elementary calculations show that $\inf\{(l-1)/|\log R(l)|\}$ is attained for $l=5$. Thus we have

THEOREM 2. *The best $\bar{\varphi}$-optimal Lehmer—Schur method in the class of congruent coverings is based on the extremal 5-disk covering of the unit disk* ([1]) *and its $\bar{\varphi}$-efficiency index is* 18.621.

REFERENCES

[1] Bezdek, K., Über einige Kreisüberdeckungen, *Beiträge Algebra Geom.* **14** (1983), 7—13. *MR* **85a**: 52012.

[2] Fejes Tóth, L., *Regular figures*, Pergamon Press, Oxford, 1964. *MR* **29** # 2705.

[3] Friedli, A., Optimal covering algorithms in methods of search for solving polynomial equations, *J. Assoc. Comput. Mach.* **20** (1973), 290—300. *MR* **49** # 1760.

[4] Henrici, P., Methods of search for solving polynomial equations, *J. Assoc. Comput. Mach.* **17** (1970), 273—283. *MR* **43** # 7058.

[5] Lehmer, D. H., A machine method for solving polynomial equations, *J. Assoc. Comput. Mach.* **2** (1961), 151—163.

[6] Lehmer, D. H., Search procedures for polynomial equation solving, *Constructive Aspects of the Fundamental Theorem of Algebra* (eds.: Dejon, B., Henrici, P.), Wiley, London, 1969. *MR* **42** # 2652.

[7] Molnár, J., Egy elemi geometriai szélsőértékfeladat, *Mat. Fiz. Lapok* **49** (1942), 249—253.

[8] Ralston, A., *A first course in numerical analysis*, McGraw-Hill, New-York, 1965. *MR* **32** # 8479.

[9] Škljarskiĭ, D. O., Čencov, N. N. and Jaglom, I. M., *Geometrical estimates and problems from combinatorial geometry*, Nauka, Moscow, 1974 (in Russian). *MR* **52** # 15205.

[10] Stewart III, G. W., On Lehmer's method for finding the zeros of a polynomial, *Math. Comp.* **23** (1969), 829—835. *MR* **42** # 1331.

(Received November 13, 1986)

AGRÁRTUDOMÁNYI EGYETEM
MATEMATIKAI INTÉZETE
SZÁMITÁSTECHNIKAI TANSZÉK
PÁTER KÁROLY U. 1
H—2103 GÖDÖLLŐ
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 461—471

ON THE STRUCTURE OF THE SOLUTIONS OF AN AUTONOMOUS DIFFERENTIAL-DELAY SYSTEM BY THE METHOD OF CHARACTERISTIC EQUATION

S. CSATÓ

1. Introduction

The asymptotic behaviour of the solutions of a linear autonomous differential-difference equation is determined by the roots of the characteristic equation. These characteristic roots can be used for establishing boundedness and stability of solutions as well as some properties of unbounded solutions. E.g., as it was proved by F. V. Atkinson and S. N. Zhang [1], the characteristic equation of the scalar equation

$$\dot{x}(t) = p\big(x(t)-x(t-r)\big) \tag{1.1}$$

has at most one root with a positive real part. From this fact it follows that every solution $x(t)$ of (1.1) is of the form

$$x(t) = y(t)c+x_0(t), \tag{1.2}$$

where y and x_0 is an unbounded and a bounded solution of (1.1), respectively; c is constant. Atkinson and Zhang proved representation (1.2) for the equation

$$\dot{x}(t) = P(t)\big(x(t)-x(t-r)\big),$$

too, where $P(t)$ is a diagonal matrix. So the question arises whether representation (1.2) holds in the case of non-diagonal $P(t)$.

In this paper we prove that the answer is negative, but a more general representation can be obtained instead of (1.2). Investigating the characteristic equation of the system

$$\dot{x}(t) = A\big(x(t)-x(t-r)\big) \tag{1.3}$$

we give conditions for the eigenvalues of A which guarantee the existence of k and only k solutions with positive real parts of the characteristic equation of (1.3). Therefore, the solutions of (1.3) can be obtained in the form

$$x(t) = Y(t)c+x_0(t),$$

Research (partially) supported by Hungarian National Foundation for Scientific Research Grant No. 6032/6319.

1980 *Mathematics Subject Classification* (1985 Revision). Primary 34K15.
Key words and phrases. Zeros of complex quasi-polynomial, representation of unbounded solutions.

where Y is a matrix whose columns are unbounded solutions of (1.3), x_0 is a bounded solution of (1.3), and c is a constant vector. It will be pointed out that while the characteristic equation of (1.1) has at most one root with a positive real part, the number of such roots of the characteristic equation of (1.3) can be an arbitrarily large natural number.

2. The results

Consider the equation

$$\dot{x}(t) = A\big(x(t)-x(t-r)\big), \tag{2.1}$$

where A is an $n\times n$ real matrix, $r>0$ is a real number. The characteristic equation of (2.1) reads

$$\det\big(A(1-e^{-rz})-zE\big) = 0. \tag{2.2}$$

If the value of the function $z/(1-e^{-rz})$ at $z=0$ is meant by its limit, equation (2.2) can be rewritten into the form

$$(1-e^{-rz})^n \det\left(A-\frac{z}{1-e^{-rz}}E\right) = 0. \tag{2.3}$$

If $\mu=z/(1-e^{-rz})$ is an eigenvalue of A then number z is a solution of this equation. This leads us to the equation

$$z-\mu(1-e^{-rz}) = 0. \tag{2.4}$$

To characterize the solutions of (2.2) we have to know the zeros of the quasi-polynomial

$$f(z) = z-\mu(1-e^{-rz}),$$

so first we investigate the localization on the complex plane of the zeros of f. If $\mu=0$, then f has the single zero $z=0$, thus we can assume $\mu\neq 0$.

Introducing the notation $\mu=Re^{i\varphi}$ $(R>0,\ -\pi<\varphi\leqq\pi)$ we get

$$f(z) = z-Re^{i\varphi}(1-e^{-rz}).$$

It is known [1] that for real μ the function f has at most one zero with a positive real part. It turns out that for a complex μ the function f may have an arbitrarily large number of zeros.

We fix those curves on the domain $R>0$, $-\pi<\varphi\leqq\pi$ of the plain $(R;\varphi)$ at whose points the real part of one of the zeros of f is zero. Since the real parts of the zeros of f are continuous functions of R and φ, these curves divide the plane in domains where the number of the zeros of f with positive real parts is constant [2].

Using the canonical form $z=x+iy$ of the complex number z in the equation $f(z)=0$ and separating the real and imaginary parts we get the systems of equations

$$x-R\cos\varphi+Re^{-rx}\cos(\varphi-ry) = 0 \tag{2.5}$$

$$y-R\sin\varphi+Re^{-rx}\sin(\varphi-ry) = 0. \tag{2.6}$$

The equation $f(z)=0$ has a solution with zero real part if y is a solution of the system

$$-R\cos\varphi+R\cos(\varphi-ry)=0, \tag{2.7}$$

$$y-R\sin\varphi+R\sin(\varphi-ry)=0. \tag{2.8}$$

By equations (2.7), (2.8), $z=0$ is always a solution of the equation $f(z)=0$, and this equation has another solution with zero real part if and only if μ is a point of the curves

$$\begin{aligned} G_0 = \{(R,\varphi)|Rr\ \sin\varphi &= \varphi, && \text{if}\quad 0<|\varphi|<\pi,\\ R &= 1/r, && \text{if}\quad \varphi=0\},\\ G_k = \{(R,\varphi)|Rr\,|\sin\varphi| &= |\varphi|+k\pi, && \text{if}\quad 0<|\varphi|<\pi\}\quad k=1,2,3,\dots. \end{aligned}$$

These curves divide the plane (R,φ) in domains

$$T_0=\Big\{(R,\varphi)|0<R<\frac{1}{r},\quad \text{if}\quad \varphi=0,\quad 0<R<\frac{\varphi}{r\sin\varphi},$$
$$\text{if}\quad 0<|\varphi|<\pi,\quad 0<R,\quad \text{if}\quad \varphi=\pi\Big\},$$

$$T_1=\Big\{(R,\varphi)|R>\frac{1}{r},\quad \text{if}\quad \varphi=0,\quad \frac{\varphi}{r\sin\varphi}<R<\frac{|\varphi|+\pi}{r|\sin\varphi|},\quad \text{if}\quad 0<|\varphi|<\pi\Big\},$$

$$T_k=\Big\{(R,\varphi)|\frac{|\varphi|+(k-1)\pi}{r|\sin\varphi|}<R<\frac{|\varphi|+k\pi}{r|\sin\varphi|},\quad \text{if}\quad 0<|\varphi|<\pi\Big\},\quad k=2,3,\dots.$$

Curves G_k and domains T_k can be seen on Figure 1.

Consider also the curve L on (R,φ) plane defined by

$$L=\Big\{(R,\varphi)|-\frac{\pi}{2}<\varphi<\frac{\pi}{2},\ \cos\varphi=\frac{1+\ln Rr}{Rr},\quad Rr\geqq 1\Big\}.$$

The following theorem describes the connection between the zeros of f and the curves and domains defined above.

THEOREM 1. (i) *f has at most double zeros.*

(ii) *f has a double zero if and only if $\mu\in L\cap G_{2k}$. The double zeros different from 0 have positive real parts.*

(iii) *If $\mu\in G_k\cup T_k$, then f has k and only k zeros with positive real parts.*

PROOF. Statement (i) follows from the formulae

$$f'(z)=1-\mu re^{-rz},\quad f''(z)=\mu r^2e^{-rz}.$$

To prove statement (ii) separate equation $f'(z)=0$ into real and imaginary part:

$$1-Rre^{-rx}\cos(\varphi-ry)=0, \tag{2.9}$$

$$Rre^{-rx}\sin(\varphi-ry)=0. \tag{2.10}$$

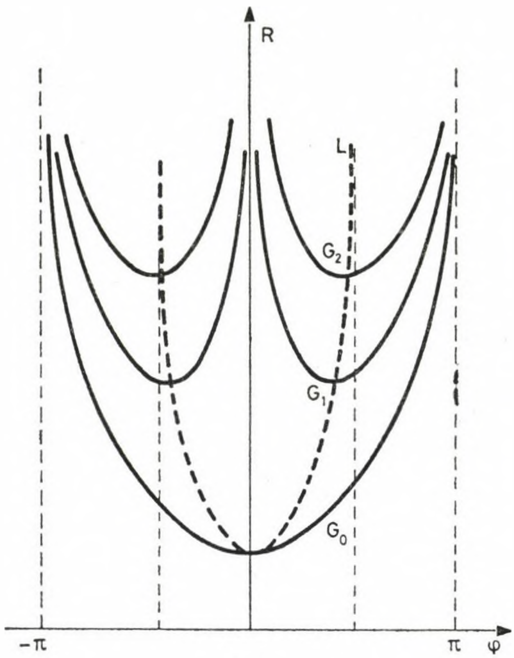

Fig. 1

The equation $f'(z)=0$ has a solution with zero real part if and only if the equations

$$1-Rr\cos(\varphi-ry)=0 \tag{2.11}$$

$$Rr\sin(\varphi-ry)=0 \tag{2.12}$$

hold. Since equation (2.7), (2.8), (2.11), (2.12) admit the only solution $y=0$, function f has the only double zero $z=0$. From equation (2.10) we get $\varphi-ry=k\pi$, where, by equation (2.9), k is even: $k=2m$ $(m=0, \pm 1, \pm 2, \ldots)$. Then equation (2.9) implies

$$1=Rre^{-rx},$$

and from (2.5) we obtain the equation

$$x=R\cos\varphi-\frac{1}{r}. \tag{2.13}$$

There exists number x satisfying both of these equations only if φ and R fulfil the equation

$$R\cos\varphi-\frac{1}{r}=\frac{\ln Rr}{r}. \tag{2.14}$$

This yields curve L. Equations (2.6) and (2.12) concern the imaginary parts of the double zeros. Number y can be a solution of these equations only if φ and R satisfy the condition

$$\varphi - 2m\pi = Rr \sin \varphi,$$

being in the definition of curve G_{2m}.

Relations (2.13) and (2.14) show also that a double zero different from 0 must have a positive real part.

In order to prove statement (iii) we show the following assertions:

(a) If $\mu \in T_0$, then f has no zero with a positive real part.

(b) If $\mu \in G_k$, then f has only one zero with zero real part besides $z=0$.

(c) If, while μ is changing, the point (R, φ) goes over to the domain T_k from T_{k-1}, then the real parts of the zeros of f with zero real parts increase.

Setting $\varphi=\pi$ in equations (2.5), (2.6) and squaring we obtain

$$(x+R)^2+y^2 = R^2 e^{-2rx}.$$

Since this equation cannot be satisfied by an $x>0$, assertion (a) is proved.

To prove (b) we observe that equations (2.7), (2.8) have only one solution different from zero. On the other hand, by statement (ii) of Theorem 1, this zero is simple.

In order to prove (c) we have to investigate how the real parts of the zeros with zero real parts vary when the point (R, φ) goes over to T_k from T_{k-1} crossing the curve G_k. To do this we need the sign of the expression

$$dx = -Re \frac{\dfrac{\partial f}{\partial R} dR + \dfrac{\partial f}{\partial \varphi} d\varphi}{\partial f / \partial z}$$

when $x=0$ and $(R, \varphi) \in G_k$ [2]. Let φ be a constant and after differentiation set $x=0$. Then a simple calculation gives

$$dx = \frac{\cos\varphi - \cos(\varphi - ry) + Rr(1-\cos ry)}{(1-Rr\cos(\varphi-ry))^2 + (Rr\sin(\varphi-ry))^2} dR.$$

Taking into account equation (2.7) and the fact that if φ is constant and (R, φ) goes over to T_k from T_{k-1}, then R increases, we obtain the inequality

$$dx = \frac{Rr(1-\cos ry)}{(1-Rr\cos(\varphi-ry))^2 + (Rr\sin(\varphi-ry))^2} dR \geq 0.$$

From this inequality it follows, too, that dx can vanish only at a certain value of y, namely on the line $\varphi=0$. If $\varphi \neq 0$, then $dx>0$; consequently, the real parts of the zeros of f with zero real parts increase when μ goes over to T_k from T_{k-1} parallel with the line $\varphi=0$ while R is increasing [2].

On the basis of assertions (a) and (b) we can describe the zeros of f in the following way. If $\mu \in T_0$, then all the zeros have negative real parts except $z=0$. When μ is moving and comes to G_0, then one of the zeros with negative real parts becomes a zero with zero real part. When μ arrives T_1, $z=0$ remains a zero, while the zero with zero real part becomes a zero with a positive real part. If $\mu \in T_1$, then f has a single zero with positive real part, $z=0$ is another zero of f, all the other zeros have negative real parts. If $\mu\in G_2$, then one of the zeros with negative real parts becomes a zero with zero real part because of $dx>0$. Therefore, if $\mu\in T_2$, then f has already two zeros with positive real parts, the other zeros have negative real parts except $z=0$.

The theorem is proved.

Let us return to equations (2.2) and (2.3). Since $\lim_{z\to 0} z(1-e^{-rz})^{-1}=1/r$, every solution of (2.2) is a solution of (2.3), too. Denote by $\mu_1, \dots, \mu_n$ the eigenvalues of matrix A and introduce the notation

$$f_i(z) = z-\mu_i(1-e^{-rz}).$$

Theorem 2 describes the connection between the zeros of f_i and the solutions of the characteristic equation (2.2).

THEOREM 2. *A complex number is a solution of equation* (2.2) *if and only if it is a zero of some function f_i. If z is a solution of* (2.2) *with multiplicity k, and it is a zero of f_i with multiplicity m_i, then $k=\sum_{i=1}^{n} m_i$.*

PROOF. First we prove the theorem for the solution $z=0$.

Number $z=0$ is a zero of every f_i and, by (2.3), it is a solution of (2.2) with a multiplicity least n. If $\mu_i\neq 1/r$, then $z=0$ is a solution of (2.3) with the (exact) multiplicity n, and then it is a simple zero of every f_i. If $\mu_i=1/r$ is an eigenvalue of A of multiplicity j, then $z=0$ is a solution of (2.3) with the multiplicity $n+j$. In this case $z=0$ is a double zero of those j functions among $f_1(z), \dots, f_n(z)$ which belong to $1/r$. Therefore, $\sum_{i=1}^{n} m_i=n+j$.

If $z\neq 0$ is a solution of (2.2) then it is a solution of the equation

$$\det\left(A-\frac{z}{1-e^{-rz}}E\right) = 0,$$

too, i.e. z is a solution of some of the equations $z(1-e^{-rz})^{-1}=\mu_i$. On the other hand, the solutions of these equations are zeros of $f_i(z)$.

The proof is complete.

Using Theorems 1 and 2 we can describe how the solutions of the characteristic equation (2.2) depend on the eigenvalues of A.

Let μ be an eigenvalue of A of multiplicity k. Denote by λ any zero of the function f belonging to μ. If $\lambda\neq 0$ is a single zero of f, then it is a solution of (2.2) of multiplicity k. If it is a double zero of f, then it is a solution of (2.2) of multiplicity $2k$. If $1/r$ is not eigenvalue of A, then $\lambda=0$ is a solution of (2.2) of multiplicity n. If $1/r$ is an eigenvalue of A with multiplicity k, then $\lambda=0$ is a solution of (2.2) of multiplicity $n+k$. Each solution of equation (2.1) belonging to λ is of the form $p(t)e^{\lambda t}$, where p is a polynomial. We distinguish two kinds of solutions: the bounded and the unbounded ones. If $\operatorname{Re}\lambda>0$, or $\operatorname{Re}\lambda=0$ and the degree of p is greater than 0, then $p(t)e^{\lambda t}$ is an unbounded solution of (2.1). If $\operatorname{Re}\lambda<0$ or $\operatorname{Re}\lambda=0$ and the degree of p is not greater than 0, then $p(t)e^{\lambda t}$ is a bounded solution of (2.1). We choose linearly independent solutions from the solutions of the form $p(t)e^{\lambda t}$.

DEFINITION. Functions $x_1, x_2, \ldots, x_k$ are linearly independent on $[a, b]$ if from identity

$$c_1x_1(t)+c_2x_2(t)+\ldots+c_kx_k(t) \equiv 0, \qquad a \leqq t \leqq b$$

it follows that $c_1=c_2=c_3=\ldots=c_k=0$.

The function $x(t)=p_j(t)e^{\lambda t}$ is a solution of equation (2.1) if the equation

$$\lambda p_j(t)+\dot{p}_j(t) = (1-e^{-\lambda r})Ap_j(t)+e^{-\lambda r}A(p_j(t)-p_j(t-r)) \tag{2.15}$$

holds.

First we investigate the solutions of equation (2.1) belonging to $\lambda=0$, e.g. the polynomial solutions. In this case the following lemma holds.

LEMMA 1. (a) *If* $1/r$ *is not eigenvalue of* A, *then the solutions of* (2.1) *belonging to* $\lambda=0$ *are constant vectors.*

(b) *If* $1/r$ *is an eigenvalue of* A *with multiplicity* $k(>0)$, *then besides the constant vectors there are still* k *linearly independent polynomials which are solutions of* (2.1), *and the degree of these polynomials is greater than* 0.

PROOF. Let $p_j(t)=v_{j1}t^j+v_{j2}t^{j-1}+\ldots+v_{jj}t+v_{jj+1}$. In this case equation (2.15) is equivalent to equation

$$\dot{p}_j(t) = A(p_j(t)-p_j(t-r)).$$

Separating this equation, we get the system of equations

$$\begin{aligned} jv_{j1} &= jrAv_{j1} \\ (j-1)v_{j2} &= (j-1)rAv_{j2}-\binom{j}{2}r^2Av_{j1} \\ &\vdots \\ v_{jj} &= rAv_{jj}-r^2Av_{jj-1}+\ldots+(-1)^{j+1}r^jAv_{j1}. \end{aligned} \tag{2.16}$$

If $1/r$ is not eigenvalue of A, then from the first equation of the system (2.16) follows that $j=0$. Consequently, in this case the solutions of equation (2.1) belonging to $\lambda=0$ are constant vectors. The statement (a) is proved.

Let us now prove case (b). Suppose that vectors $h_1, h_2, \ldots, h_j$ exist such that $Ah_1=1/r\, h_1$, $Ah_i=1/r\, h_i+h_{i-1}$, $i=2, 3, \ldots, j$. It is easy to see that the system (2.16) has a unique system of solutions $v_{j1}, v_{j2}, \ldots, v_{jj}$ which can be expressed by the vectors $h_1, h_2, \ldots, h_j$ as

$$v_{ji} = \alpha_{j1}^{(i)}h_1+\alpha_{j2}^{(i)}h_2+\ldots+\alpha_{ji}^{(i)}h_i,$$

where α_{jk} are constant, $\alpha_{ji}^{(i)}\neq 0$ $(i=1, 2, \ldots, j)$. Hence statement (b) follows.

Let us now consider the solutions of (2.1) belonging to $\lambda\neq 0$.

LEMMA 2. *Suppose that* μ *is an eigenvalue of* A *of multiplicity* k.

(a) *If* λ *is a single zero of* f, *then there are polynomials* $p_i(t)$ $(i=1, \ldots, k)$ *such that the degree of* $p_i(t)$ *is not greater than* $k-1$ *and the functions* $p_i(t)e^{\lambda t}$ *are linearly* *independent solutions of* (2.1), *moreover* $p_i(t_0)e^{\lambda t_0}$ *are linearly independent vectors for every* $t_0 \geqq 0$.

(b) *If* λ *is a double zero of* f, *then there are* $2k$ *linearly independent solutions of* (2.1) *of the form* $p_i(t)e^{\lambda t}$ $(i=1, 2, \dots, 2k)$ *such that the degree of* $p_i(t)$ *is not greater than* $2k-1$.

PROOF. If $\lambda \neq 0$, then the equation (2.15) can be transformed into the form

$$(A-\mu E)p_j(t) = \frac{\dot{p}_j(t)}{1-e^{-\lambda r}} - \frac{e^{-\lambda r}}{1-e^{-\lambda r}} A\big(p_j(t)-p_j(t-r)\big).$$

Separating this equality we get the system of the following equations

$$\begin{aligned}
&(A-\mu E)v_{j1} = 0\\
&(A-\mu E)v_{j2} = j\,\frac{v_{j1}-re^{-\lambda r}Av_{j1}}{1-e^{-\lambda r}}\\
(2.17)\qquad &(A-\mu E)v_{j3} = (j-1)\,\frac{v_{j2}-re^{-\lambda r}Av_{j1}}{1-e^{-\lambda r}} + \frac{r^2 e^{-\lambda r}}{1-e^{-\lambda r}}\binom{j}{2}Av_{j1}\\
&\qquad\vdots\\
&(A-\mu E)v_{jj+1} = \frac{v_{jj}-re^{-\lambda r}Av_{jj}}{1-e^{-\lambda r}} + \frac{r^2 e^{-\lambda r}Av_{jj-1}}{1-e^{-\lambda r}} + \dots + \frac{(-1)^j r^j e^{-\lambda r}Av_{j1}}{1-e^{-\lambda r}}.
\end{aligned}$$

Suppose that the vectors $h_1, h_2, \dots, h_{j+1}$ exist, such that $Ah_1=\mu h_1$, $Ah_i=\mu h_i+h_{i-1}$, $i=2, \dots, j, j+1$. It is easy to see that in this case the system (2.17) has a unique system of solutions $v_{j1}, v_{j2}, \dots, v_{jj+1}$ which can be expressed by the vectors h_i as

$$v_{ji} = \alpha_{j1}^{(i)} h_1 + \alpha_{j2}^{(i)} h_2 + \dots + \alpha_{ji}^{(i)} h_i, \qquad i = 1, 2, \dots, j+1,$$

where $\alpha_{jk}^{(i)}$ are constant, $\alpha_{ji}^{(i)} \neq 0$ $(k=1, 2, \dots, j+1)$. Therefore it follows that the vectors v_{ji} $(i=1, 2, \dots, j+1)$ are linearly independent, that is the statement (a) is true.

(b) If λ is a double zero of function f, then $1-\mu re^{-\lambda r}=0$ and equations (2.17) are transformed to the system of equations

$$\begin{aligned}
&(A-\mu E)v_{j1} = 0\\
&(A-\mu E)v_{j2} = 0\\
(2.18)\qquad &(A-\mu E)v_{j3} = -(j-1)\frac{re^{-\lambda r}}{1-e^{-\lambda r}}u_1 + \frac{r^2 e^{-\lambda r}}{1-e^{-\lambda r}}\binom{j}{2}Av_{j1}\\
&\qquad\vdots\\
&(A-\mu E)v_{jj+1} = -\frac{re^{-\lambda r}}{1-e^{-\lambda r}}u_{j-1} + \frac{r^2 e^{-\lambda r}}{1-e^{-\lambda r}}Av_{jj-1} + \dots + \frac{(-1)^j r^j e^{-\lambda r}}{1-e^{-\lambda r}}Av_{j1},
\end{aligned}$$

where $u_i=Av_{ji+1}-\mu v_{ji+1}$, $i=1, 2, \dots, j-1$. If to the eigenvalue μ belong the vectors $h_1, h_2, \dots, h_m$ such that $Ah_1=\mu h_1$, $Ah_i=\mu h_i+h_{i-1}$ $(i=2, \dots, m)$ and $j=2m-1$.

then the system (2.18) has a unique system of solutions

$$v_{2m-11}, v_{2m-12}, \ldots, v_{2m-12m},$$

which can be expressed by the vectors h_i as

$$v_{2m-12i-1} = \alpha^{(i)}_{2m-11} h_1 + \alpha^{(i)}_{2m-12} h_2 + \ldots + \alpha^{(i)}_{2m-1i} h_i$$
$$v_{2m-12i} = \beta^{(i)}_{2m-11} h_1 + \beta^{(i)}_{2m-12} h_2 + \ldots + \beta^{(i)}_{2m-1i} h_i,$$

$i=1, 2, \ldots, m$, where $\alpha^{(i)}_{2m-1i}$, $\beta^{(i)}_{2m-1i}$ $(i=1, 2, \ldots, m)$, are constant, $\alpha^{(i)}_{2m-1i}\neq 0$ and $\beta^{(i)}_{2m-1i}\neq 0$. The Lemma 2 is proved.

Let us denote the number of the unbounded linearly independent solutions of equation (2.1) by m. Let $\mu_1, \mu_2, \ldots, \mu_j$ be the eigenvalues of A (different from $1/r$ and 0), let us denote their multiplicity by $k_1, k_2, \ldots, k_j$. Suppose that $\mu_1 \in T_{i_1}, \mu_2 \in T_{i_2}, \ldots, \mu_N \in T_{i_N}$ $(0 \leq N \leq j)$, $\mu_{N+1} \in G_{i_{N+1}}, \mu_{N+2} \in G_{i_{N+2}}, \ldots, \mu_j \in G_{i_j}$. Let

$$k_+ = k_1 i_1 + k_2 i_2 + \ldots + k_N i_N,$$
$$k_0 = k_{N+1}(i_{N+1}+1) - l_{N+1} + k_{N+2}(i_{N+2}+1) - l_{N+2} + \ldots + k_j(i_j+1) - l_j,$$

where l_i is the number of the eigenvectors belonging to μ_i, $N+1 \leq i \leq j$. Let us denote the multiplicity of the eigenvalue $1/r$ by k_r. Then $m = k_+ + k_0 + k_r$.

Now we can formulate the following theorem on the solutions of equation (2.1).

THEOREM 3. *Every solution x of equation* (2.1) *can be represented in the form*

$$x(t) = Y(t)c + x_0(t), \tag{2.19}$$

where Y is an $n \times m$ matrix whose columns are unbounded linearly independent solutions of equation (2.1), *$c \in \mathbf{C}^m$ is a constant vector and x_0 is a bounded solution of equation* (2.1).

This theorem generalizes Theorems 1.1, 1.2 and 1.3 of F. V. Atkinson and S. N. Zhang [1] concerned with the scalar equation

$$\dot{x}(t) = p\big(x(t) - x(t-r)\big). \tag{2.20}$$

Namely, it follows from our Theorem 1 that in the case $p < 1/r$ the solutions of the characteristic equation of (2.20) have negative real parts except the solution $z=0$; therefore in this case every solution of (2.20) is of the form

$$x(t) = x_0(t).$$

If $p=1/r$, then $z=0$ is a characteristic root of multiplicity 2, the other characteristic roots have negative real parts; therefore, in this case every solution of (2.20) is of the form

$$x(t) = at + x_0(t).$$

If $p > 1/r$, then the characteristic equation of (2.20) has a positive real solution denote it by λ. Then the representation

$$x(t) = ae^{\lambda t} + x_0(t)$$

follows from our Theorem 3.

If all the eigenvalues of matrix A lie in domain T_0, then each solution of (2.1) has a limit as $t\to\infty$ [5]. Integrating equation (2.1) we can calculate this limit, so we get the following assertion:

COROLLARY 2.1. *If* $\mu\in T_0$, *then each solution of* (2.1) *has a limit as* $t\to\infty$. *The limit of the solution through* $(0, g)$ *equals*

$$(E-rA)^{-1}g(0)-\int_{-r}^{0}Ag(t)\,dt.$$

Lemmas 1, 2 and Theorem 3 imply the concept of asymptotically ordinary functional differential equations introduced and investigated by I. Győry [3, 4].

A linear functional differential equation is called asymptotically ordinary on the interval $[t_0, \infty)$ if there is a function $u\colon [t_0-r, \infty)\to \mathbf{R}^n\times\mathbf{R}^n$ with the following properties

(i) $\det(u(t))\neq 0$ for all $t\geqq t_0-r$,

(ii) for every solution x the relation

$$x(t) = u(t)\big(c(x)-o(1)\big),\quad t\to\infty$$

holds,

(iii) for every $c\in\mathbf{R}^n$ there is a solution x such that $c(x)=c$.

COROLLARY 2.2. *If all the eigenvalues of A lie in the domain T_1, then equation* (2.1) *is asymptotically ordinary.*

PROOF. If eigenvalue μ_i lies in the domain T_1, then f_i has one zero with positive real parts, e.g. Y is an $n\times n$ matrix. If the multiplicity of μ_i equals k_i then the columns of Y belonging to the eigenvalue μ_i are of the form $p_{ij}(t)e^{\lambda_i t}$ $(j=1, 2, \dots, k_i)$, where λ_i is zero of f_i with positive real parts, $p_{ij}(t)$ is a polynomial (its degree is not greater than k_i-1). From statement (a) of Lemma 2 it follows that the vectors $p_{ij}(t_0)e^{\lambda_i t_0}$ are linearly independent for every $t_0\geqq 0$. Consequently, the columns of $Y(t_0)$ are linearly independent vectors for every $t_0\geqq 0$, e.g. $\det(Y(t))\neq 0$. In this case $Y^{-1}(t)$ exists and since $\operatorname{Re}\lambda_i>0$ it follows that

$$\lim_{t\to\infty}\|Y^{-1}(t)\| = 0.$$

If matrix Y is real, then let $u(t)\equiv Y(t)$ and according to the formulae (2.19) the Corollary 2.2 is true. Let us consider the case where Y is not real. Let us denote the j-th column of Y by Y_j, that is $Y=(Y_1, Y_2, \dots, Y_n)$. If Y_j is real then let us define $u_j(t)\equiv Y_j(t)$, where u_j is the j-th column of u. If Y_j is not real, then there is a k such that $Y_k=\overline{Y}_j$. ($\overline{Y}_j$ is the complex conjugate of Y_j.) In this case we can defined the j-th and k-th columns of u in the following way: $u_j(t)\equiv 1/2\,(Y_j(t)+Y_k(t))$, $u_k(t)\equiv 1/2i\,(Y_j(t)-Y_k(t))$. Since $\operatorname{Re} Y_j(t_0)$ and $\operatorname{Im} Y_j(t_0)$ are linearly independent vectors for every $t_0\geqq 0$, then $\det(u(t))\neq 0$ for every $t\geqq 0$, $u^{-1}(t)$ exists and

$$\lim_{t\to\infty}\|u^{-1}(t)\| = 0.$$

By the formulae (2.19) the Corollary 2.2 is proved.

ACKNOWLEDGEMENT. The author would like to thank Professor J. Terjéki for helpful suggestions.

REFERENCES

[1] ATKINSON, F. V. and ZHANG, S. N., Asymptotic behaviour and structure of solutions for equation $\dot{x}(t)=p(t)(x(t)-x(t-r))$ (to appear).

[2] EL'SGOL'S, L. E. and NORKIN, S. B., *Introduction to the theory of differential equations with deviating argument,* Nauka, Moscow, 1971 (in Russian). *MR* **50** # 5133.

[3] GYŐRI, I., On asymptotically ordinary functional differential equations, *Differential Equations,* (Colloq., Keszthely, 1974), Coll. Math. Soc. János Bolyai **15**, North-Holland, Amsterdam, 1977. *MR* **58** # 1468.

[4] GYŐRI, I., On existence of the limits of solutions of functional-differential equations, *Qualitative Theory of Differential Equations,* Vol. I, II (Szeged, 1979), Colloq. Math. Soc. J. Bolyai, 30, North-Holland, Amsterdam—New York, 1981. *MR* **84**g: 34130.

[5] HALE, J. K., *Functional differential equations,* Applied Mathematical Sciences, Vol. 3, Springer-Verlag, New York—Heidelberg—Berlin, 1971. *Zbl* **222** # 34003; *MR* **57** # 6711.

(Received February 9, 1987)

ÉLELMISZERIPARI FŐISKOLA
GÉPÉSZETI ÉS AUTOMATIZÁLÁSI INTÉZET
MARX TÉR 7
H—6724 SZEGED
HUNGARY

Studia Scientiarum Mathematicarum Hungarica 24 (1989), 473—481

ON THE GLIVENKO—CANTELLI THEOREM FOR BALLS IN METRIC SPACES

TAMÁS SZABADOS

Abstract

It is shown that the Glivenko—Cantelli theorem holds for continuous measures over the class of balls with centers in a fixed compact set or over the class of balls in a finite dimensional space, while it does not hold in an infinite dimensional real inner product space.

1. Introduction

Let $(S, \mathscr{A}, P)$ denote a probability space and let $(S^n, \mathscr{A}^n, P^n)$ be its n-th product $(1 \leqq n \leqq \infty)$. If $Z=(Z_1, Z_2, \ldots) \in S^\infty$ is a random sample, then

$$\mu_n^Z = \frac{1}{n} \sum_{i=1}^{n} \delta_{Z_i}$$

denotes the corresponding empirical measure $(n=1, 2, \ldots)$, where δ_{Z_i} is the unit mass concentrated at the point Z_i. If $\mathscr{V}$ is a subclass of $\mathscr{A}$, then the $\mathscr{V}$-discrepancy between the measures P and Q is defined as

$$D(P, Q; \mathscr{V}) = \sup_{A \in \mathscr{V}} |P(A) - Q(A)|.$$

$\mathscr{V}$ is called a Glivenko—Cantelli (GC) class for P if

$$\lim_{n \to \infty} D(P, \mu_n^Z; \mathscr{V}) = 0 \quad P^\infty \text{ a.e.}$$

In the sequel S is a metric space and $\mathscr{A}$ is taken as the Borel σ-field in S (i.e. the smallest σ-field containing all open subsets of S). The main topic of this paper is to study the GC property of the class $\mathscr{B}$ of all closed balls in S. Particularly, when S is a vector space as well, this property is closely related to the GC property of the class $\mathscr{H}$ of all closed half-spaces in S. It turned out that there is an essential difference between the finite an infinite dimensional cases. (A good survey about the evolution of the GC problem can be found in Gaenssler and Stute (1979).)

Fortet and Mourier (1953) showed that in a k-dimensional Euclidean space the set $\mathscr{H}$ is a GC class for any P which is absolutely continuous w.r.t. Lebesgue measure. Wolfowitz (1954, 1960) generalized this result to arbitrary P's.

1980 *Mathematics Subject Classification* (1985 Revision). Primary 62G30; Secondary 60B10.

Key words and phrases. Glivenko—Cantelli class, empirical measure, probability in metric spaces.

Elker (1975) proved that in R^k $\mathcal{B}$ is a GC class for arbitrary P's. This also follows from a general theorem of Vapnik and Chervonenkis (1971) combining their result with Dudley (1979). Elker, Pollard and Stute (1979) gave a rather elementary method, based on the so-called university classes and compactness arguments, to prove that (among other classes of convex sets) $\mathcal{H}$ and $\mathcal{B}$ are GC classes in R^k for arbitrary P's.

On the other hand, when S is an infinite dimensional space, there are mostly negative answers to the GC problem. Sazonov (1963) showed that $\mathcal{H}$ is not a GC class in this situation in general. In fact, it follows from his result that in an infinite dimensional real vector space, $\mathcal{H}$ is not a GC class for any *$\mathcal{H}$-continuous* measure P, i.e. if P assigns zero measure to the boundary of any half-space. Also, Tøpsoe, Dudley and Hoffmann—Jørgensen (1976) gave an example that $\mathcal{B}$ is not a GC class in an infinite dimensional Banach space in general.

In this paper it will be shown that $\mathcal{B}_K$, the class of closed balls with centers in an arbitrary compact subset K, is a GC class in any metric space S if P is *$\mathcal{B}$-continuous,* i.e. if P assigns zero measure to the boundary of any ball. Then, in case of $\mathcal{B}$-continuous P's, this result plus simple geometric arguments will be applied to obtain an elementary proof of the (already known) fact that $\mathcal{B}$ is a GC class in any finite dimensional real inner product space. Finally, Sazonov's (1963) result combined with simple geometric consideration will be applied to prove that $\mathcal{B}$ is not a GC class in an infinite dimensional real inner product space whenever P is $\mathcal{H}$-continuous.

R. M. Dudley kindly called my attention to references [2] and [9]. Talagrand (1984) gives a general solution to the GC problem, so my results should follow from his. Also, Dudley (1984) describes a rather general method for determining GC convergence using the notion of total boundedness in the sense of metric entropy with inclusion. This could be also applied to prove my results. Of course, the results of this paper should follow from the general theory of Vapnik and Chervonenkis (1971) as well.

2. The compact case

We need the following elementary fact about ordering. (Obviously, this has been known for a long time, but as I cannot give a reference, the proof will be also included.)

LEMMA 1. *Let $x_1, \dots, x_n$ and $y_1, \dots, y_n$ be real numbers. Let us denote by $x_{(1)} \leqq \dots \leqq x_{(n)}$ and by $y_{(1)} \leqq \dots \leqq y_{(n)}$, respectively, the above real numbers in non-decreasing order. Then for any convex function ψ we have*

$$\sum_{i=1}^{n} \psi(x_{(i)} - y_{(i)}) \leqq \sum_{i=1}^{n} \psi(x_i - y_i).$$

PROOF. Since one can order numbers by performing a sequence of transpositions, it is enough to prove that

$$\psi(a-c) + \psi(b-d) \leqq \psi(a-d) + \psi(b-c),$$

if $a \leqq b$ and $c \leqq d$. But the convexity of ψ implies that $\psi(x-c)-\psi(x-d)$ $(c \leqq d)$ is non-decreasing as a function of x, so that

$$\psi(a-c)-\psi(a-d) \leqq \psi(b-c)-\psi(b-d). \quad \blacksquare$$

COROLLARY. *For any* $p \geqq 1$,

$$\Big[\sum_{i=1}^{n} |x_{(i)}-y_{(i)}|^p\Big]^{1/p} \leqq \Big[\sum_{i=1}^{n} |x_i-y_i|^p\Big]^{1/p},$$

and if $p \to \infty$ *we get*

$$\max_{1 \leqq i \leqq n} |x_{(i)}-y_{(i)}| \leqq \max_{1 \leqq i \leqq n} |x_i-y_i|. \tag{1}$$

THEOREM 1. *Let K be a compact subset of the metric space S and let P be a $\mathscr{B}$-continuous probability measure in S. Then $\mathscr{B}_K$, the set of all balls with centers in K, is a GC class for P.*

PROOF. For any $Z=(Z_1, Z_2, \dots) \in S^\infty$ and $n=1, 2, \dots$ we take the empirical measure μ_n^Z and we define

$$d_n^Z(x) = \sup_{r>0} \big|P(B(x, r))-\mu_n^Z(B(x, r))\big| \quad (x \in S), \tag{2}$$

where $B(x, r)$ denotes the closed ball with centre at x and radius r. If ϱ denotes the metric in S, then $d_n^Z(x)$ is clearly a Kolmogorov statistics for the random variable $r_j^Z(x)=\varrho(x, Z_j)$ for any fixed $x \in S$, and

$$d_n^Z(x) = \max_{1 \leqq j \leqq n;\, k=0,1} \left|P(B(x, r_{(j)}^Z(x)))-\frac{j-k}{n}\right|, \tag{3}$$

where $r_{(1)}^Z(x) \leqq \dots \leqq r_{(n)}^Z(x)$ are the distances $r_1^Z(x), \dots, r_n^Z(x)$ in non-decreasing order.

Let us take any $\varepsilon>0$. First, the classical GC theorem implies that for any $x \in S$ there exists a set $W(x) \subset S^\infty$ of P^∞-measures zero and for any $Z \notin W(x)$ there is an integer $N(x, Z)$ such that

$$d_n^Z(x) < \varepsilon \quad \text{if} \quad Z \notin W(x) \quad \text{and} \quad n \geqq N(x, Z). \tag{4}$$

Second, we prove that d_n^Z is continuous on S, uniformly in Z and n. I.e., for each $x \in S$ there exists a $\delta(x)>0$ such that for any $Z \in S^\infty$ and $n=1, 2, \dots$ we have

$$|d_n^Z(x)-d_n^Z(y)| < \varepsilon \quad \text{if} \quad \varrho(x, y) < \delta(x). \tag{5}$$

To show this, by (3), it is enough to prove that

$$\big|P(B(x, r_{(j)}^Z(x)))-P(B(y, r_{(j)}^Z(y)))\big| < \varepsilon \quad \text{if} \quad \varrho(x, y) < \delta(x)$$

for any $Z \in S^\infty$, $n=1, 2, \dots,$ and $j=1, \dots, n$. Since P is $\mathscr{B}$-continuous, this follows if we show that for any $Z \in S^\infty$, $n=1, 2, \dots,$ and $j=1, \dots, n$ we have

$$|r_{(j)}^Z(x)-r_{(j)}^Z(y)| < \delta(x) \quad \text{if} \quad \varrho(x, y) < \delta(x).$$

But this is a consequence of (1) with $x_j=r_j^Z(x)=\varrho(x, Z_j)$ and $y_j=r_j^Z(y)=\varrho(y, Z_j)$, $j=1, \dots, n$.

Finally, since K is compact, we have points $t_1, \dots, t_m$ in K such that

$$K \subset \bigcup_{i=1}^{m} B(t_1, \delta(t_i)),$$

where $\delta(t)$ is defined according to (5). Let

$$W = \bigcup_{i=1}^{m} W(t_i) \quad \text{and} \quad N(Z) = \max_{1\leq i\leq m} N(t_i, Z) \quad \text{if} \quad Z \notin W.$$

For any $y\in K$ we can find a point t_i so that $\varrho(t_i, y)<\delta(t_i)$, therefore, by (4) and (5), we obtain

$$d_n^Z(y) < 2\varepsilon \quad \text{if} \quad Z\notin W \quad \text{and} \quad n \geqq N(Z).$$

This proves that

$$\lim_{n\to\infty} D(P, \mu_n^Z; \mathscr{B}_K) = \lim_{n\to\infty} \sup_{x\in K} d_n^Z(x) = 0 \qquad P^\infty \text{ a.e.} \quad \blacksquare$$

3. Two geometric lemmas

We need the following simple geometric fact, which essentially says that a ball with a distant centre and a large radius is similar to a half space. In a real inner product space S a closed half-space $H(x, \alpha)$ will be defined as

$$H(x, \alpha) = \{z\colon \langle z, x\rangle \geqq \alpha\} \quad (x\in S, \alpha\in R^1).$$

We also introduce the notation U for the unit sphere in S:

$$U = \{z\colon |z| = 1\} = \{z\colon \langle z, z\rangle = 1\}.$$

LEMMA 2. *Let S be a real inner product space and let us fix a closed ball $B_0=B(0, R_0)$ with an arbitrary radius $R_0>0$. Then for any $x_0\in U$ and $\delta>0$ there exist numbers $R_1=R_1(x_0, \delta, R_0)>R_0$ and $\eta=\eta(x_0, \delta, R_0)>0$ such that*

$$y\in S, \quad |y| \geqq R_1, \quad \left|\frac{y}{|y|}-x_0\right| \leqq \eta \tag{6}$$

imply

$$B_0\cap H(x_0, \alpha_2) \subset B_0\cap B(y, r) \subset B_0 \cap H(x_0, \alpha_1)$$

for every $r>0$ with suitable numbers α_1 and α_2, $0<\alpha_2-\alpha_1<\delta$.

PROOF. First, for any $z\in B_0\cap B(y, r)$ we will choose an α_1 such that $z\in B_0\cap H(x_0, \alpha_1)$. If

$$|z-y|^2 = |z|^2+|y|^2-2|z||y|\left\langle\frac{z}{|z|}, \frac{y}{|y|}-x_0\right\rangle-2|y|\langle z, x_0\rangle \leqq r^2, \tag{7}$$

then

$$\langle z, x_0\rangle \geqq \frac{1}{2|y|}\left(|z|^2+|y|^2-r^2-2|z||y|\left\langle\frac{z}{|z|}, \frac{y}{|y|}-x_0\right\rangle\right).$$

If $|z|\leqq R_0$ also holds, then choosing

$$\alpha_1 = \frac{1}{2|y|}(|y|^2-r^2-2|y|R_0\eta),$$

by (6) we have $\langle z, x_0\rangle \geqq \alpha_1$, i.e. $z\in H(x_0, \alpha_1)$.

In a similar manner, if $z\in B_0\cap H(x_0, \alpha_2)$ and so $\langle z, x_0\rangle\geqq\alpha_2$ holds, where

$$\alpha_2 = \frac{1}{2|y|}(R_0^2+|y|^2-r^2+2|y|R_0\eta),$$

then by (7) we obtain that $z\in B_0\cap B(y, r)$.

Since now

$$0 < \alpha_2-\alpha_1 < \frac{1}{2|y|}(R_0^2+4|y|R_0\eta) \leqq R_0^2/(2R_1)+2R_0\eta,$$

the lemma is proved. ∎

Lemma 2 is closely related to the next one, which states that in any real inner product space $\mathscr{B}$-continuity implies $\mathscr{H}$-continuity. (Its converse is not true; an example: uniform distribution along the unit circle in R^2.)

Lemma 3. *Let S be a real inner product space and let P be a $\mathscr{B}$-continuous probability measure in S. Then the P-measure of any hyperplane $\{z\colon \langle z, x_0\rangle=\alpha_0\}$ is zero, i.e. P is $\mathscr{H}$-continuous as well.*

Proof. Let $B_0=B(0, R_0)$ be a closed ball in S with arbitrary radius $R_0>0$. Then we show that for any $x_0\in U$, $\alpha_0\in R$ and $\delta>0$ there exist $y\in S$ and numbers r_1, r_2, $0<r_2-r_1<\delta$, such that

$$B_0\cap\{z\colon \langle z, x_0\rangle = \alpha_0\} \subset B_0\cap[B(y, r_2)\setminus B(y, r_1)]. \tag{8}$$

For, let us take $y=Rx_0$ with a suitably chosen $R>0$, and let $z\in B_0$, $\langle z, x_0\rangle= =\alpha_0$. Then

$$|z-y|^2 = |z|^2+R^2-2R\alpha_0, \quad |z| \leqq R_0.$$

We choose

$$r_2 = (R_0^2+R^2-2R\alpha_0)^{1/2}, \quad r_1 = (R^2-2R\alpha_0)^{1/2},$$

since then (8) follows and

$$0 < r_2-r_1 = R_0^2/(r_1+r_2) \to 0 \quad \text{as} \quad R\to\infty. \ ∎$$

4. The finite dimensional case

Our aim here is to give a new, elementary proof to the following, already known fact.

THEOREM 2. *Let S be a finite dimensional real inner product space and let P be a $\mathscr{B}$-continuous probability measure in S. Then $\mathscr{B}$ is a GC class for P.*

PROOF. We have to show that

$$\lim_{n\to\infty} D(P, \mu_n^Z, \mathscr{B}) = \lim_{n\to\infty} \sup_{y\in S} \sup_{r>0} |P(B(y, r)) - \mu_n^Z(B(y, r))| = 0 \quad P^\infty \text{ a.e.}$$

Let us fix an $\varepsilon>0$. Then there exists a closed ball $B_0=B(0, R_0)$ such that

$$P(B_0^c) < \varepsilon/5. \tag{9}$$

Then by the *SLLN* there exists a set $W_0\subset S^\infty$ of P^∞-measure zero and for any $Z\notin W_0$ there exists an integer $N_0(Z)$ such that we have

$$\mu_0^Z(B_0^c) < 2\varepsilon/5 \quad \text{if} \quad n \geqq N_0(Z). \tag{10}$$

Since P is $\mathscr{H}$-continuous by Lemma 3, for any $x_0\in U$ one can choose a $\delta(x_0)>0$ such that

$$0 \leqq P(H(x_0, \alpha_1)) - P(H(x_0, \alpha_2)) < \varepsilon/5 \tag{11}$$

if $0<\alpha_2-\alpha_1<\delta(x_0)$. Then, by Lemma 2, there exist numbers $R_1(x_0)>R_0$ and $\eta(x_0)>0$ such that $y\in S$, $|y|\geqq R_1(x_0)$, $\big|y/|y|-x_0\big|\leqq\eta(x_0)$ imply

$$B_0\cap H(x_0, \alpha_2) \subset B_0\cap B(y, r) \subset B_0\cap H(x_0, \alpha_1) \tag{12}$$

for every $r>0$ with suitable numbers α_1 and α_2, $0<\alpha_2-\alpha_1<\delta(x_0)$.

Because of the compactness of the unit sphere U, we have points $t_1, \dots, t_m$ on U such that U is covered by the balls $B(t_i, \eta(t_i))$, $i=1, \dots, m$. Then we take

$$R_1 = \max_{1\leqq i\leqq m} R_1(t_i).$$

By Theorem 1, for the compact ball $B_1=B(0, R_1)$ we have

$$\lim_{n\to\infty} \sup_{y\in B_1} \sup_{r>0} |P(B(y, r)) - \mu_n^Z(B(y, r))| = 0 \quad P^\infty \text{ a.e.} \tag{13}$$

In order to handle the balls with centers in B_1^c we will consider half-spaces in the directions $t_1, \dots, t_m$. According to the classical GC theorem, for each t_i there is a set $W_i\subset S^\infty$ of P^∞-measure zero and for any $Z\notin W_i$ there is an integer $N_i(Z)$ such that

$$\sup_{\alpha\in R^1} |P(H(t_i, \alpha)) - \mu_n^Z(H(t_i, \alpha))| < \varepsilon/5 \quad \text{if} \quad n \geqq N_i(Z). \tag{14}$$

Now we choose

$$W = \bigcup_{i=0}^{m} W_i \quad \text{and} \quad N(Z) = \max_{0\leqq i\leqq m} N_i(Z) \quad \text{if} \quad Z\notin W.$$

If $y \in B_1^c$ then

$$\left|\frac{y}{|y|}-t_i\right| < \eta(t_i)$$

for some t_i. By (9), (11), and (12), for any $r>0$ we obtain α_1 and α_2 such that

$$0 \leqq P(H(t_i, \alpha_1))-P(H(t_i, \alpha_2)) < \varepsilon/5 \tag{15}$$

and

$$P(H(t_i, \alpha_2))-\varepsilon/5 < P(B(y, r)) < P(H(t_i, \alpha_1))+\varepsilon/5. \tag{16}$$

Also, by (10) and (12), for arbitrary $Z \notin W$ and $n \geqq N(Z)$ we get

$$\mu_n^Z(H(t_i, \alpha_2))-2\varepsilon/5 < \mu_n^Z(B(y, r)) < \mu_n^Z(H(t_i, \alpha_1))+2\varepsilon/5. \tag{17}$$

Finally, it follows from (14), ..., (17) that

$$\sup_{y \in B_1^c} \sup_{r>0} \left|P(B(y, r))-\mu_n^Z(B(y, r))\right| < \varepsilon$$

if $Z \notin W$ and $n \geqq N(Z)$. This and (13) prove the theorem. ▮

5. The infinite dimensional case

THEOREM 3. *Let S be an infinite dimensional real inner product space and let P be an $\mathcal{H}$-continuous probability measure in S. Then for some* $\varepsilon>0$,

$$\liminf_{n \to \infty} D(P, \mu_n^Z; \mathcal{B}) \geqq \varepsilon \quad P^\infty \text{ a.e.}$$

PROOF. By Theorem 1 in Sazonov (1963), the assumptions above imply that there exists an $\varepsilon>0$ such that $P^\infty(E)=1$, where

$$E = \bigcap_{n=1}^{\infty} \left\{Z \in S^\infty \colon \sup_{|x|=1} \left|P(H(x, 0))-\mu_n^Z(H(x, 0))\right| \geqq 6\varepsilon\right\}. \tag{18}$$

Let $B_0=B(0, R_0)$ be a closed ball in S with the property

$$P(B_0^c) < \varepsilon, \tag{19}$$

and let $E_0 \subset S^\infty$ be a set of P^∞-measure one and $N_0(Z)$ be an integer for each $Z \in R_0$ such that

$$\mu_n^Z(B_0^c) < 2\varepsilon \quad \text{if} \quad n \geqq N_0(Z). \tag{20}$$

For any $Z \in E \cap E_0$ and for any $n \geqq N_0(Z)$ we choose a point $x_0=x_0(\varepsilon, Z, n)$ on the unit sphere U with the property

$$\left|P(H(x_0, 0))-\mu_n^Z(H(x_0, 0))\right| \geqq 5\varepsilon. \tag{21}$$

Then we take a $\delta(x_0)>0$ such that

$$0 \leqq P(H(x_0, \alpha_1))-P(H(x_0, \alpha_2)) < \varepsilon \quad \text{if} \quad 0 \leqq \alpha_2-\alpha_1 < \delta(x_0). \tag{22}$$

By Lemma 2, there exists a number $R_1(x_0) > R_0$ such that taking $y = R_2 x_0$, $R_2 \geqq R_1(x_0)$ and $r = (R_2^2 + R_0^2)^{1/2}$ we obtain

$$B_0 \cap H(x_0, 0) \subset B_0 \cap B(y, r) \subset B_0 \cap H(x_0, \alpha_1) \quad \text{if} \quad \alpha_1 = -R_0^2/(2R_2).$$

Here R_2 is chosen so that $|\alpha_1| < \delta(x_0)$ and

$$\mu_n^Z(H(x_0, \alpha_1)) = \mu_n^Z(H(x_0, 0)). \tag{23}$$

Now, by (19),

$$P(H(x_0, 0)) - \varepsilon < P(B(y, r)) < P(H(x_0, \alpha_1)) + \varepsilon$$

and by (20),

$$\mu_n^Z(H(x_0, 0)) - 2\varepsilon < \mu_n^Z(B(y, r)) < \mu_n^Z(H(x_0, \alpha_1)) + 2\varepsilon.$$

Hence, by (22),

$$\left|P(B(y, r)) - P(H(x_0, 0))\right| < 2\varepsilon,$$

and by (23),

$$\left|\mu_n^Z(B(y, r)) - \mu_n^Z(H(x_0, 0))\right| < 2\varepsilon.$$

These formulae and (21) show that

$$\left|P(B(y, r)) - \mu_n^Z(B(y, r))\right| > 5\varepsilon - 4\varepsilon = \varepsilon. \quad \blacksquare$$

Note that if one substitutes $\mathscr{B}$ by the class of those balls that are centered at sample points, then the GC theorem can hold even in the infinite dimensional case. More exactly, it follows from [8], Theorem 1, that

$$\lim_{n\to\infty} \max_{1\leqq i\leqq n} \sup_{r>0} \left|P(B(Z_i, r)) - \mu_n^Z(B(Z_i, r))\right| = 0 \quad P^\infty \text{ a.e.}$$

in any separable metric space S if P is $\mathscr{B}$-continuous.

Also, it is clear from Theorem 1 that if one substitutes $D(P, \mu_n^Z; \mathscr{B})$ by

$$D_Q^\delta(P, \mu_n^Z) = \sup\left\{a \in R^1\colon\ Q(\{x\colon\ d_n^Z(x) > a\}) > \delta\right\},$$

where d_n^Z is defined by (2), then

$$\lim_{n\to\infty} D_Q^\delta(P, \mu_n^Z) = 0 \quad P^\infty \text{ a.e.}$$

for each $\delta > 0$ in any metric space S if P is a $\mathscr{B}$-continuous and Q is a tight probability measure in S.

Acknowledgement. I am indebted to G. Tusnády for his valuable help and advice. I also want to thank R. M. Dudley for some corrections and suggestions.

REFERENCES

[1] Dudley, R. M., Balls in $\mathbf{R}^k$ do not cut all subsets of $k+2$ points, *Adv. in Math.* **31** (1979), 306—308. *MR* **81g**: 52010.

[2] Dudley, R. M., A course on empirical processes, *École d'été de probabilités de Saint-Flour XII—1982*, Lecture Notes in Math. **1097**, Springer, Berlin, 1984, 1—142. *MR* **88e**: 60029.

[3] ELKER, J., Über ein gleichmässiges Gesetz der grossen Zahlen, Diplomarbeit, Institut für Mathematik, Ruhr-Universität Bochum, 1975.
[4] ELKER, J., POLLARD, D. and STUTE, W., Glivenko—Cantelli theorems for classes of convex sets, *Adv. in Appl. Probab.* **11** (1979), 820—833. *MR* **81a**: 60042.
[5] FORTET, R. and MOURIER, E., Convergence de la répartition empirique vers la répartition théorique, *Ann. Sci. École Norm. Sup.* (3) **70** (1953), 267—285. *MR* **15**—808.
[6] GAENSSLER, P. and STUTE, W., Empirical processes: A survey of results for independent and identically distributed random variables, *Ann. Probab.* **7** (1979), 193—243. *MR* **80d**: 60002.
[7] SAZONOV, V. V., On the theorem of Glivenko and Cantelli, *Teor. Verojatnost. i Primenen.* **8** (1963), 299—303. *MR* **27** # 4257.
[8] SZABADOS, T., Goodness of fit tests in metric spaces based on balls around the sample, *Statist. Decisions.* **5** (1987), 381—389. *MR* **88k**: 62080.
[9] TALAGRAND, M., Solution du problème de Glivenko—Cantelli, *C. R. Acad. Sci. Paris Sér. I, Math.* **298** (1984), 213—216. *MR* **85c**: 60041.
[10] TOPSØE, F., DUDLEY, R. M. and HOFFMANN-JØRGENSEN, J., Two examples concerning uniform convergence w.r.t. balls in Banach spaces, *Empirical distributions and processes* (Selected papers, Meeting on math. stochastics, Oberwolfach, 1976), Lecture Notes in Math. **566**, Springer, Berlin, 1976, 141—146. *MR* **55** # 11327.
[11] VAPNIK, V. N. and CERVONENKIS, A. JA., The uniform convergence of frequencies of the appearance of events to their probabilities, *Teor. Verojatnost. i Primenen.* **16** (1971), 264—279. *MR* **44** # 6018.
[12] WOLFOWITZ, J., Generalization of the theorem of Glivenko—Cantelli, *Ann. Math. Statist.* **25** (1954), 131—138. *MR* **15**—808.
[13] WOLFOWITZ, J., Convergence of the empiric distribution function on half-spaces, *Contributions to Probability and Statistics*, Stanford Univ. Press, Stanford, Calif., 1960, 504—507. *MR* **22** # 11420.

(Received February 13, 1987)

BUDAPESTI MŰSZAKI EGYETEM
VILLAMOSMÉRNÖKI KAR
MATEMATIKA TANSZÉK
SZTOCZEK U. 4. H. ÉPÜLET
H—1521 BUDAPEST
HUNGARY

PRINTED IN HUNGARY

Szegedi Nyomda, Szeged

CONTENTS

HU ISSN 0081-6906

Index: 26.748